

CASE STUDY RETAIL

-Louie Antony

Requirement 1: Data Loading

The screenshot displays the SQL Server Enterprise Manager interface. The central pane shows a SQL query window with the following code:

```
-- Create the Sales Transactions Data table
CREATE TABLE SalesTransactions (
    TransactionID INT PRIMARY KEY,
    ProductID INT,
    StoreID INT,
    TransactionDate DATE,
    QuantitySold INT,
    Revenue DECIMAL(10, 2)
);

-- Insert sample data into Sales Transactions Data
INSERT INTO SalesTransactions (TransactionID, ProductID, StoreID, TransactionDate, QuantitySold, Revenue)
VALUES
(1, 101, 1, '2023-01-05', 50, 500.00),
(2, 102, 2, '2023-01-10', 40, 600.00),
(3, 103, 1, '2023-02-15', 30, 300.00),
(4, 104, 3, '2023-03-20', 60, 900.00),
(5, 101, 2, '2023-04-25', 70, 700.00),
(6, 102, 3, '2023-05-30', 45, 675.00),
(7, 105, 1, '2023-06-05', 55, 550.00),
(8, 106, 2, '2023-07-10', 38, 570.00),
(9, 107, 3, '2023-08-15', 42, 630.00),
(10, 105, 1, '2023-09-20', 68, 680.00);

-- Create the Product Data table
CREATE TABLE Products (
```

The Messages pane at the bottom shows the execution results:

```
(10 row(s) affected)
(7 row(s) affected)
(3 row(s) affected)
```

The status bar at the bottom indicates: "Query executed successfully. EC2AMAZ-OIIIOMA\SQLEXPRESS - EC2AMAZ-OIIIOMA\Admini... master 00:00:00 0 rows".

The Properties pane on the right shows connection details for "EC2AMAZ-OIIIOMA\SQLEXPRESS":

- Connection name: EC2AMAZ-OIIIOMA\SQLEXPRESS
- Connection Details:
 - Connection elapsed: 00:00:00.555
 - Connection finish time: 9/15/2023 9:16:49 AM
 - Connection rows returned: 0
 - Connection start time: 9/15/2023 9:16:49 AM
 - Connection state: Open
- Display name: EC2AMAZ-OIIIOMA\SQLEXPRESS
- Login name: EC2AMAZ-OIIIOMA\Administrator
- Server name: EC2AMAZ-OIIIOMA\SQLEXPRESS
- Server version: 12.0.2000
- Session Tracing ID: SPID 52

File Home Insert Model View Data

Clipboard: Paste, Cut, Copy, Format painter, Get data, E...

Navigator

Display Options

- EC2AMAZ-OIIIOMA\SQLEXPRESS [4]
 - PetroCorpDB
 - ReportServer\$SQLEXPRESS
 - ReportServer\$SQLEXPRESSTempDB
 - Sales [4]
 - Customers
 - Products
 - SalesTransactions
 - Stores

Select Related Tables

Stores

StoreID	StoreName	Location	StoreManager
1	Downtown	New York	John Smith
2	Uptown	Los Angeles	Jane Doe
3	Suburbia	Chicago	Mike Johnson

Visualizations

Build visual

Values

Add data fields here

Drill through

Cross-report ☐ Off

Keep all filters ☒ On

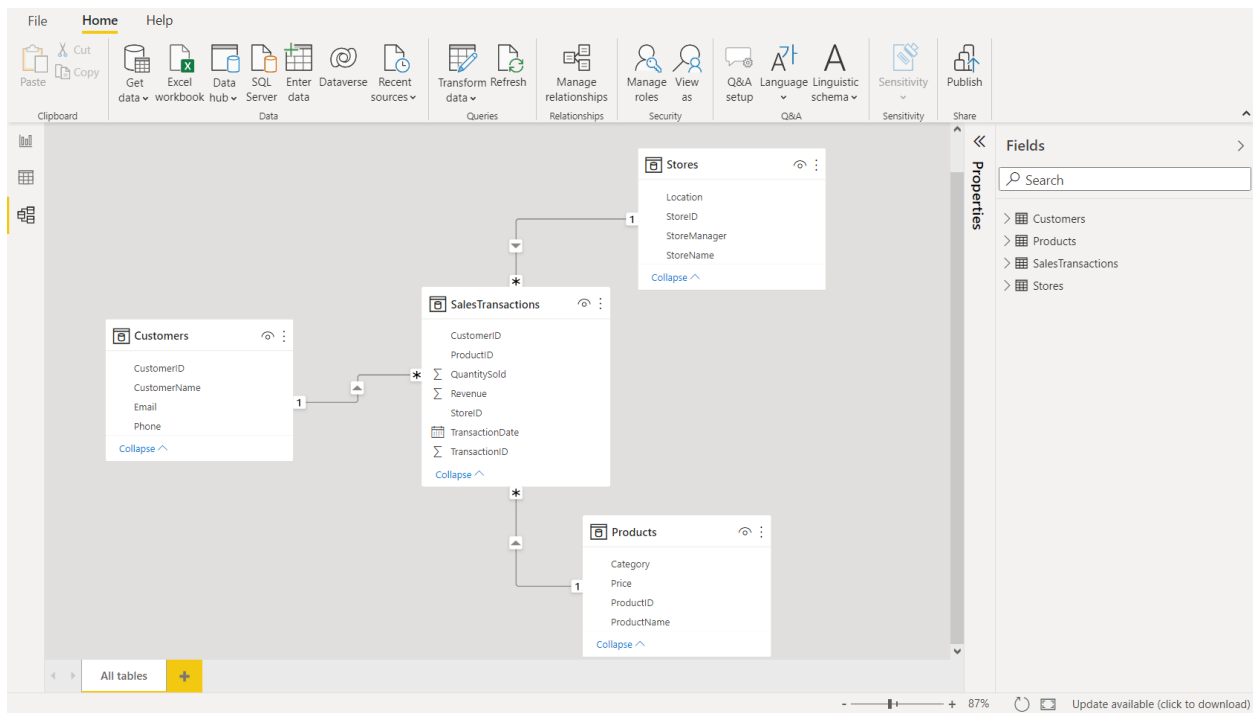
Add drill-through fields here

Page 1 of 1

80%

Update available (click to download)

9:25 AM 9/15/2023



Requirement 2: Data Transformation

Merge Tables:

The screenshot displays the Microsoft Power BI Desktop interface. The main workspace shows a data table with the following content:

ProductID	Sum of QuantitySold	StoreID	ProductName
101	50	1	Laptop XYZ
101	70	2	Laptop XYZ
102	40	2	T-Shirt Blue
102	45	3	T-Shirt Blue
103	30	1	Smartphone ABC
104	60	3	Sofa Set
105	123	1	Tablet PQR
106	38	2	Jeans Black
107	42	3	Chair
Total	498		

The right-hand pane contains the 'Fields' task pane, which is used to build visualizations. It shows a search bar and a list of data sources: Customers, Products, SalesTransactions, and Stores. The 'Products' table is expanded, showing fields: Category, Price, ProductID, and ProductName. The 'SalesTransactions' table is also expanded, showing fields: ProductID, QuantitySold, Revenue, and StoreID. The 'Columns' section of the Fields pane lists the fields currently in the visualization: ProductID, Sum of QuantitySold, StoreID, and ProductName. The 'Filters' pane on the left shows filters for ProductID, ProductName, StoreID, and Sum of QuantitySold, all set to 'is (All)'. The 'Visualizations' pane on the right shows a bar chart visualization. The 'Drill through' section is set to 'Off', and the 'Keep all filters' option is set to 'On'. The status bar at the bottom indicates 'Page 1 of 1' and '100%' zoom.

File Home Transform Add Column View Tools Help

Close & Apply New Source Recent Enter Data Data source settings Manage Parameters Data Sources Parameters

Queries [4] Customers Products SalesTransactions Stores

Merge

Select a table and matching columns to create a merged table.

Products

ProductID	ProductName	Category	Price
101	Laptop XYZ	Electronics	800
102	T-shirt Blue	Clothing	15
103	Smartphone ABC	Electronics	400
104	Sofa Set	Furniture	1200
105	Tablet PQR	Electronics	300

SalesTransactions

TransactionID	ProductID	StoreID	TransactionDate	QuantitySold	Revenue
1	101	1	1/5/2023	50	500
2	102	2	1/10/2023	40	600
3	103	1	2/15/2023	30	300
4	104	3	3/20/2023	60	900
5	101	2	4/25/2023	70	700

Join Kind
Left Outer (all from first, matching from second)

☒ Use fuzzy matching to perform the merge

> Fuzzy matching options

⚠ We were unable to determine how many matches the selection will return.

OK Cancel

4 COLUMNS, 7 ROWS Column profiling based on top 1000 rows

PREVIEW DOWNLOADED AT 9:26 AM

File Home Transform Add Column View Tools Help

Close & Apply New Source Recent Enter Data Data source settings Manage Parameters Data Sources Parameters Query Refresh Preview Advanced Editor Choose Remove Columns Remove Rows Keep Rows Remove Rows Split Column Group By Use First Row as Headers Replace Values Merge Queries Append Queries Combine Files Combine Text Analytics Vision Azure Machine Learning AI Insights

Queries [4] Customers Products SalesTransactions Stores

= Table.ExpandTableColumn(#"Merged Queries", "Customers", {"CustomerID", "CustomerName", "Email", "Phone"},

Revenue	Customers.CustomerID	Customers.CustomerName	Customers.Email	Customers.Phone
50	500	1001 Customer A	customerA@email.com	(123) 456-7890
40	600	1002 Customer B	customerB@email.com	(234) 567-8901
30	300	1003 Customer C	customerC@email.com	(345) 678-9012
60	900	1004 Customer D	customerD@email.com	(456) 789-0123
70	700	1001 Customer A	customerA@email.com	(123) 456-7890
45	675	1002 Customer B	customerB@email.com	(234) 567-8901
55	550	1005 Customer E	customerE@email.com	(567) 890-1234
38	570	1006 Customer F	customerF@email.com	(678) 901-2345
42	630	1007 Customer G	customerG@email.com	(789) 012-3456
68	680	1005 Customer E	customerE@email.com	(567) 890-1234

11 COLUMNS, 10 ROWS Column profiling based on top 1000 rows

PREVIEW DOWNLOADED AT 9:58 AM

Clean Data:

File Home Transform Add Column View Tools Help

Group By Use First Row as Headers Count Rows

Reverse Rows

Data Type: Whole Number

Replace Values

Unpivot Columns

Split Column

Format

Parse

Merge Columns

Statistics

Standard Scientific

Number Column

Trigonometry

Rounding

Information

Date

Time

Duration

Run R script

Run Python script

Scripts

Queries [4]

Customers

Products

SalesTransactions

Stores

CustomerID

Email

Phone

1 customerA@email.com (123) 456-7890

2 customerB@email.com (234) 567-8901

3 customerC@email.com (345) 678-9012

4 customerD@email.com (456) 789-0123

5 customerE@email.com (567) 890-1234

6 customerF@email.com (678) 901-2345

7 customerG@email.com (789) 012-3456

Copy

Remove

Remove Other Columns

Duplicate Column

Add Column From Examples...

Remove Duplicates

Remove Errors

Change Type

Transform

Replace Values...

Replace Errors...

Group By...

Fill

Unpivot Columns

Unpivot Other Columns

Unpivot Only Selected Columns

Rename...

Move

Drill Down

Add as New Query

Query Settings

PROPERTIES

Name

Customers

APPLIED STEPS

Source

Navigation

Removed Duplicates

4 COLUMNS, 7 ROWS Column profiling based on top 1000 rows

PREVIEW DOWNLOADED AT 9:53 AM

Create Calculated Columns:

File Home Help Table tools

Name SalesTransactions

Structure

Mark as date table

Calendars

TransactionID

ProductID

StoreID

CustomerID

1 101 1

2 102 2

3 103 1

4 104 3

5 101 2

6 102 3

7 105 1

8 106 2

9 107 3

10 105 1

Quick measure

Calculation

Multiplication

Calculate the product of two values. [Learn more](#)

Base value

Sum of QuantitySold

Value to multiply with

Sum of Price

Fields

CustomerID

CustomerName

Email

Phone

Products

Category

Price

ProductID

ProductName

SalesTransactions

CustomerID

Measure

ProductID

Products.Category

Products.Price

Products.ProductName

QuantitySold

Revenue

StoreID

TransactionDate

TransactionID

Stores

OK

Cancel

Post an idea

Table: SalesTransactions (10 rows) Column: Revenue (10 distinct values)

Update available (click to download)

File Home Help Table tools Measure tools

Name: QuantitySold x Sum of Price =

Home table: SalesTransactions

Format: \$% Format: Auto

Data category:

Structure: 1 QuantitySold x Sum of Price =
2 SUM('SalesTransactions'[QuantitySold]) * SUM(Products'[Price])

Properties: A single value for column 'QuantitySold' in table 'SalesTransactions' cannot be determined. This can happen when a measure formula refers to a column that contains many values without specifying an aggregate function.

Fields: Search

- Customers
- Products
 - Category
 - Price
 - ProductID
 - ProductName
- SalesTransactions
 - CustomerID
 - Measure
 - ProductID
 - Products.Category
 - Products.Price
 - Products.ProductNa...
 - QuantitySold
 - QuantitySol... (selected)
 - Revenue
 - StoreID
 - TransactionDate
 - TransactionID
- Stores

TransactionID	ProductID	StoreID	CustomerID	TransactionDate	QuantitySold	Revenue	Products.ProductName	Products.Category	Products.Price
1	101	1	1001	Thursday, January 5, 2023	50	500	Laptop XYZ	Electronics	800
2	102	2	1002	Tuesday, January 10, 2023	40	600	T-Shirt Blue	Clothing	15
3	103	1	1003	Wednesday, February 15, 2023	30	300	Smartphone ABC	Electronics	400
4	104	3	1004	Monday, March 20, 2023	60	900	Sofa Set	Furniture	1200
5	101	2	1001	Tuesday, April 25, 2023	70	700	Laptop XYZ	Electronics	800
6	102	3	1002	Tuesday, May 30, 2023	45	675	T-Shirt Blue	Clothing	15
7	105	1	1005	Monday, June 5, 2023	55	550	Tablet PQR	Electronics	300
8	106	2	1006	Monday, July 10, 2023	38	570	Jeans Black	Clothing	40
9	107	3	1007	Tuesday, August 15, 2023	42	630	Chair	Furniture	80
10	105	1	1005	Wednesday, September 20, 2023	68	680	Tablet PQR	Electronics	300

Table: SalesTransactions (10 rows) Column: QuantitySold x Sum of Price (0 distinct values)

Update available (click to download)

Filter Data:

File Home Help Table tools Column tools

Name: ProductID

Data type: Whole number

Format: Whole number

Summarization: Don't summarize

Data category: Uncategorized

Structure: ProductID

Sort: Sort by column

Groups: Data groups

Relationships: Manage relationships

Calculations: New column

Fields: Search

- Customers
- Products
 - Category
 - Price
 - ProductID
 - ProductName
- SalesTransactions
 - CustomerID
 - Measure
 - ProductID (selected)
 - Products.Category
 - Products.Price
 - Products.ProductNa...
 - QuantitySold
 - Revenue
 - StoreID
 - TransactionDate
 - TransactionID
- Stores

Number filters

Show rows where 'ProductID'...

is greater than 105

And Or

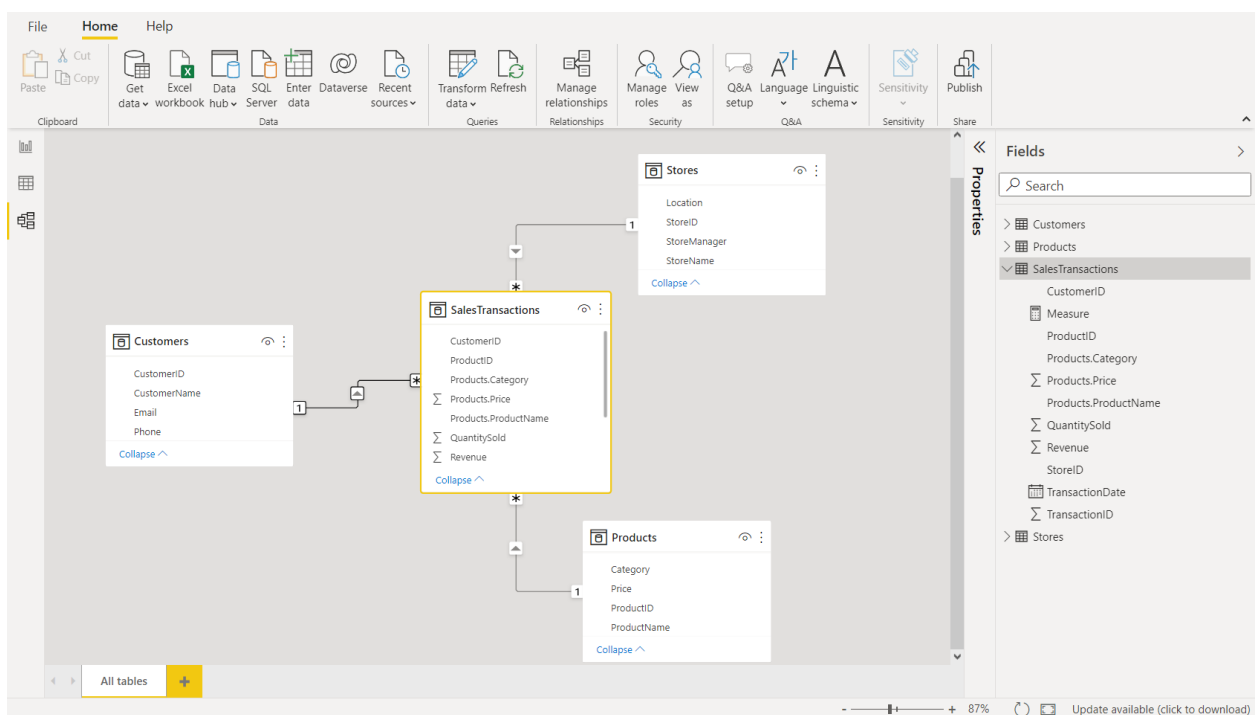
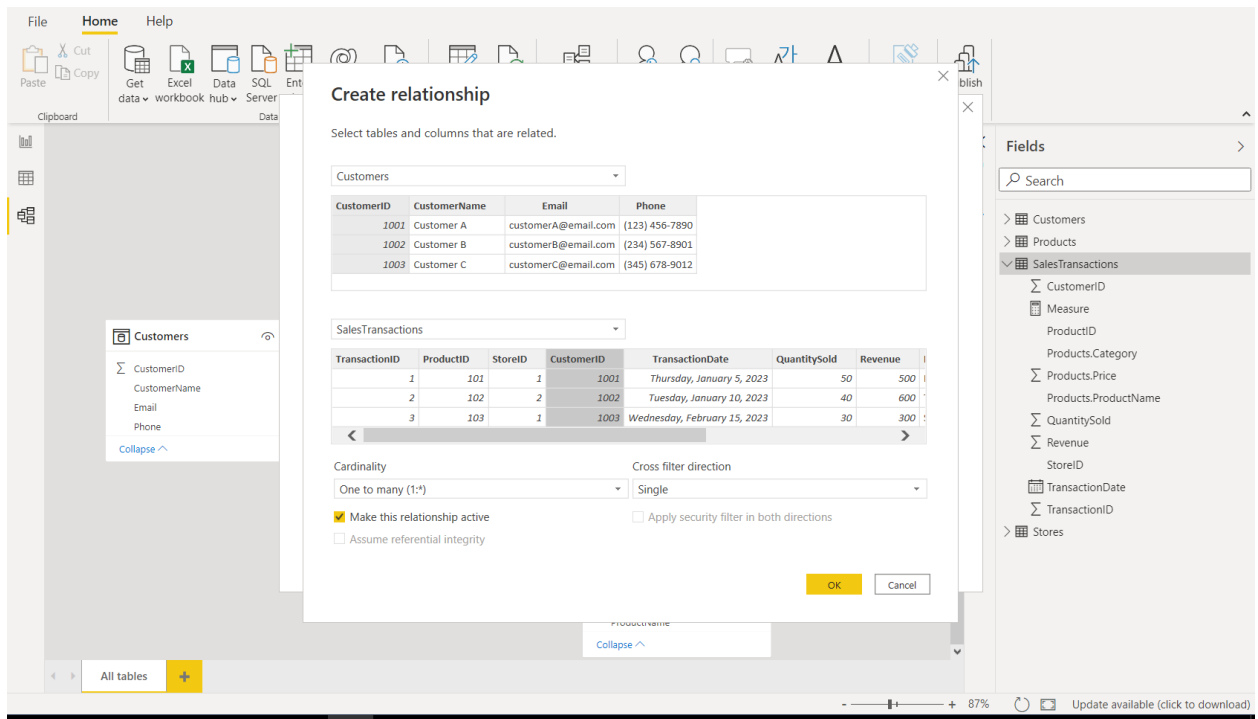
is less than 107

OK Cancel

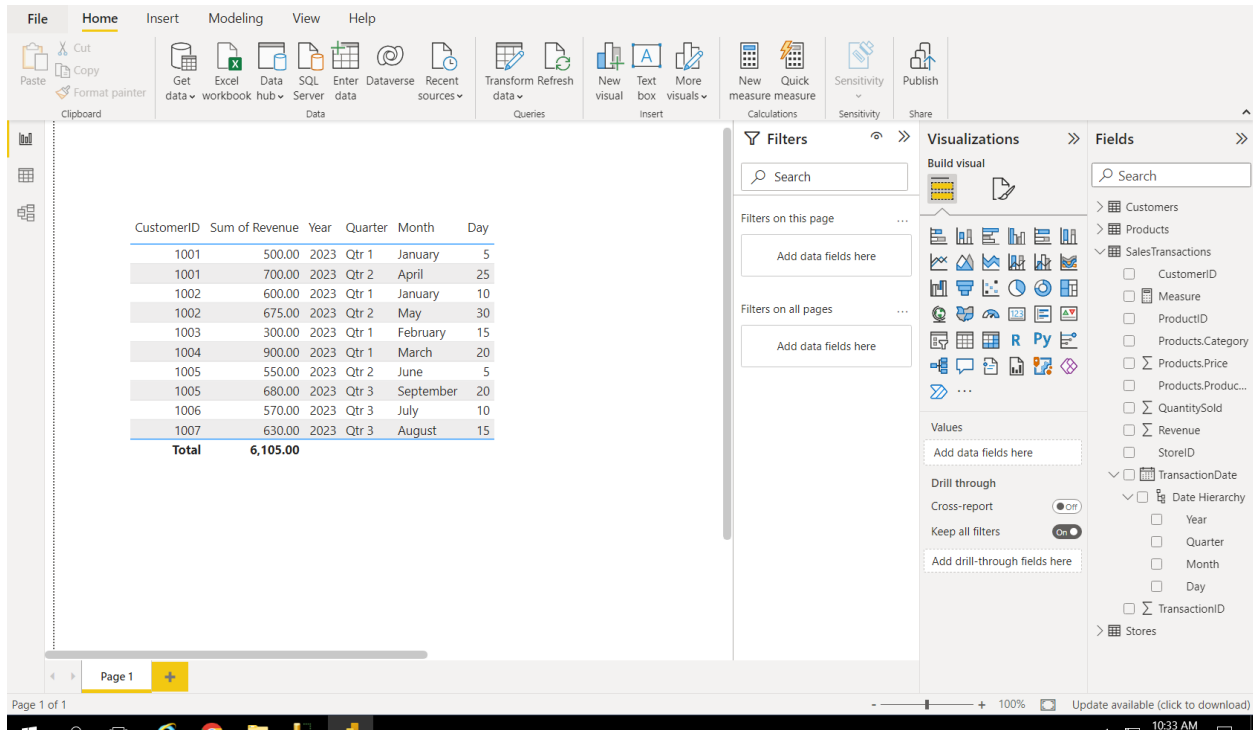
TransactionID	ProductID	StoreID	CustomerID	TransactionDate	QuantitySold	Revenue	Products.ProductName	Products.Category	Products.Price
1	101	1	1001	Thursday, January 5, 2023	50	500	Laptop XYZ	Electronics	800
2	102	2	1002	Tuesday, January 10, 2023	40	600	T-Shirt Blue	Clothing	15
3	103	1	1003	Wednesday, February 15, 2023	30	300	Smartphone ABC	Electronics	400
4	104	3	1004	Monday, March 20, 2023	60	900	Sofa Set	Furniture	1200
5	101	2	1001	Tuesday, April 25, 2023	70	700	Laptop XYZ	Electronics	800
6	102	3	1002	Tuesday, May 30, 2023	45	675	T-Shirt Blue	Clothing	15
7	105	1	1005	Monday, June 5, 2023	55	550	Tablet PQR	Electronics	300
8	106	2	1006	Monday, July 10, 2023	38	570	Jeans Black	Clothing	40
9	107	3	1007	Tuesday, August 15, 2023	42	630	Chair	Furniture	80
10	105	1	1005	Wednesday, September 20, 2023	68	680	Tablet PQR	Electronics	300

Table: SalesTransactions (10 rows) Column: ProductID (7 distinct values)

Update available (click to download)

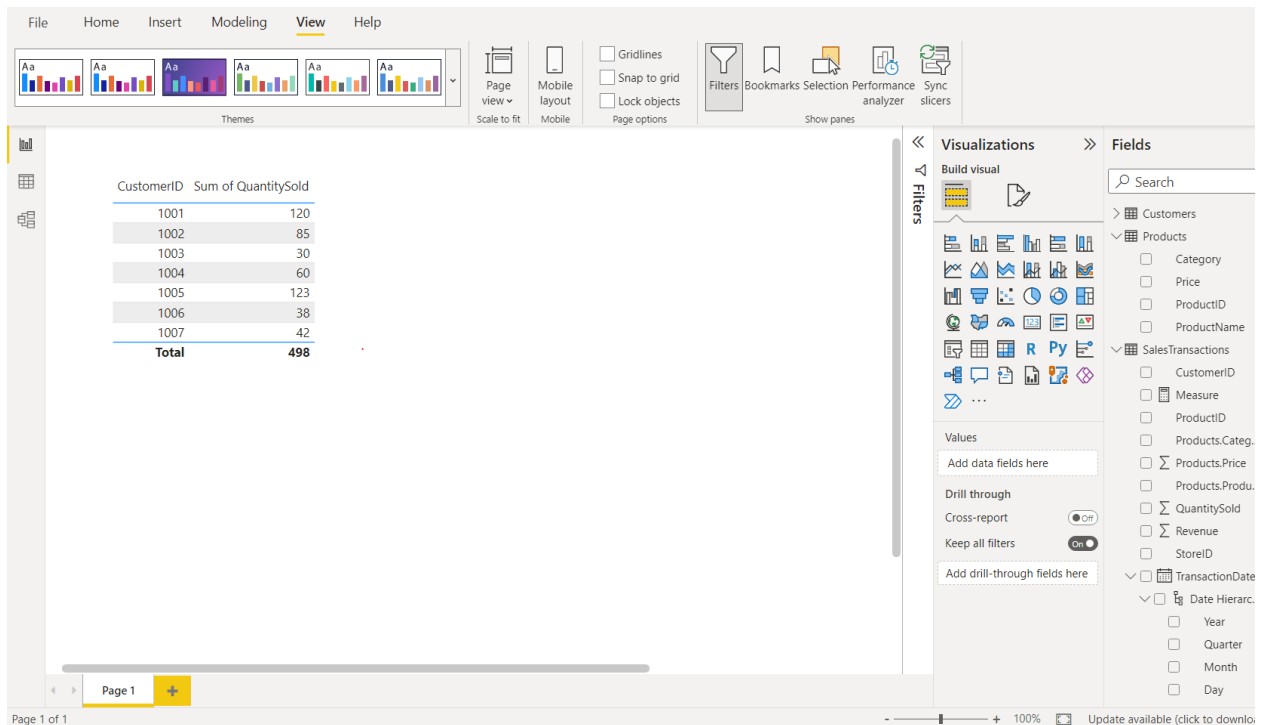


Create Hierarchies:

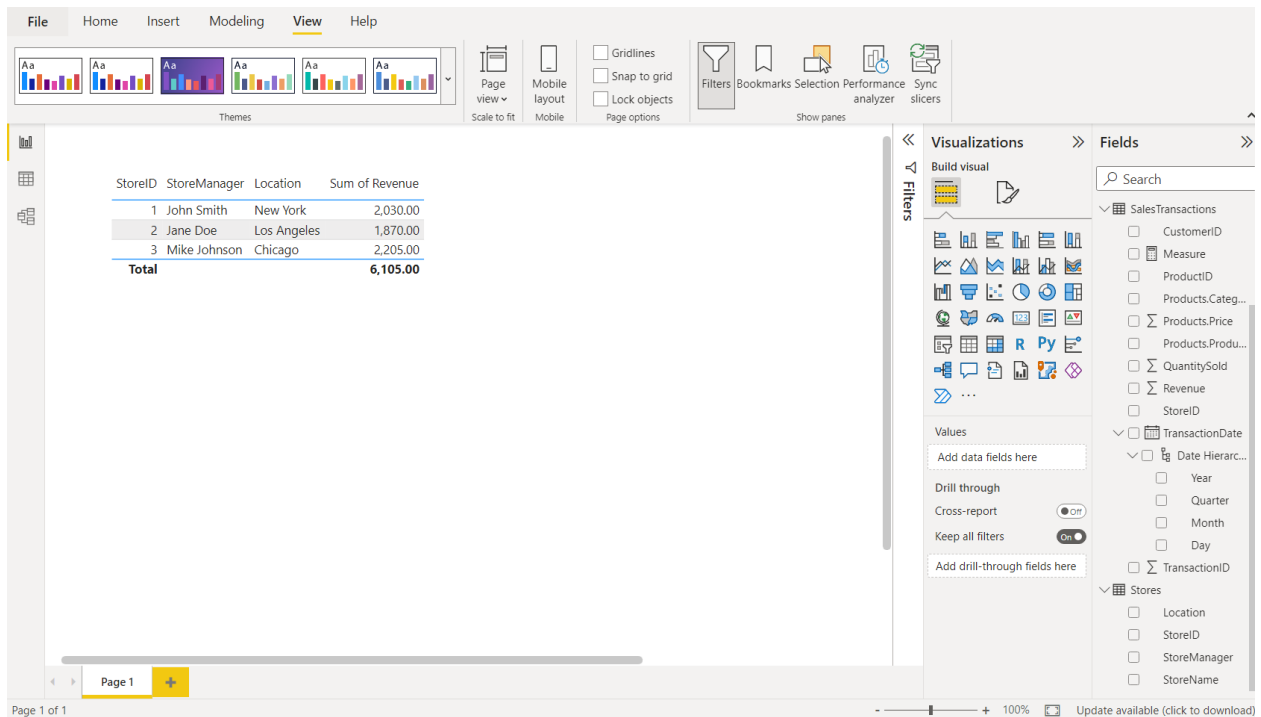


Requirement 4: Business Queries and Analysis

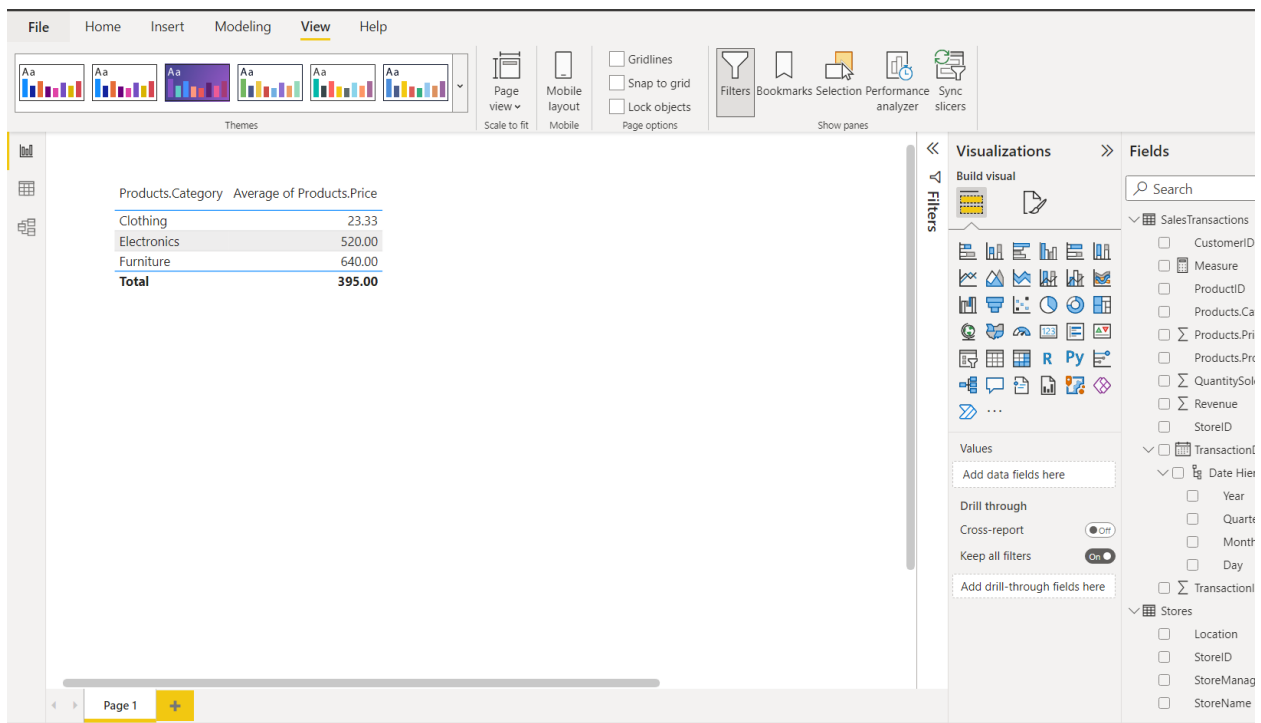
Who are the top-spending customers based on their total purchase amount?



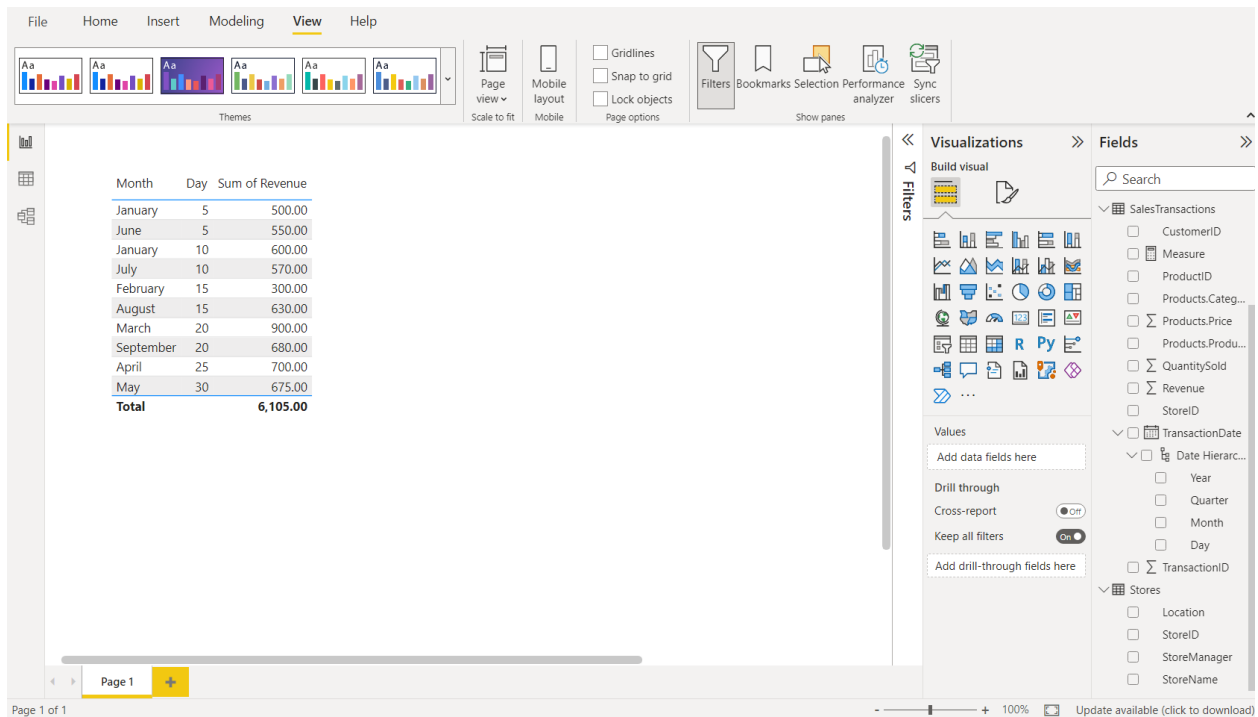
How is sales revenue distributed among different store managers?



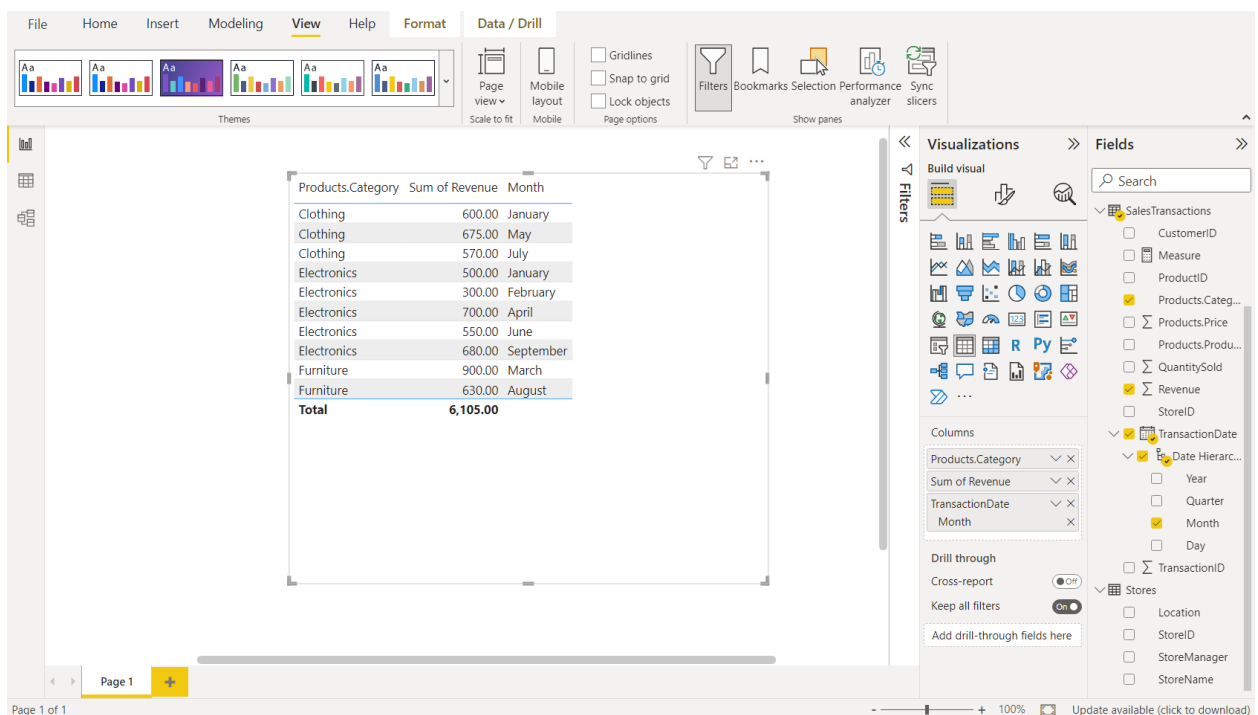
What is the average price of products in each category?



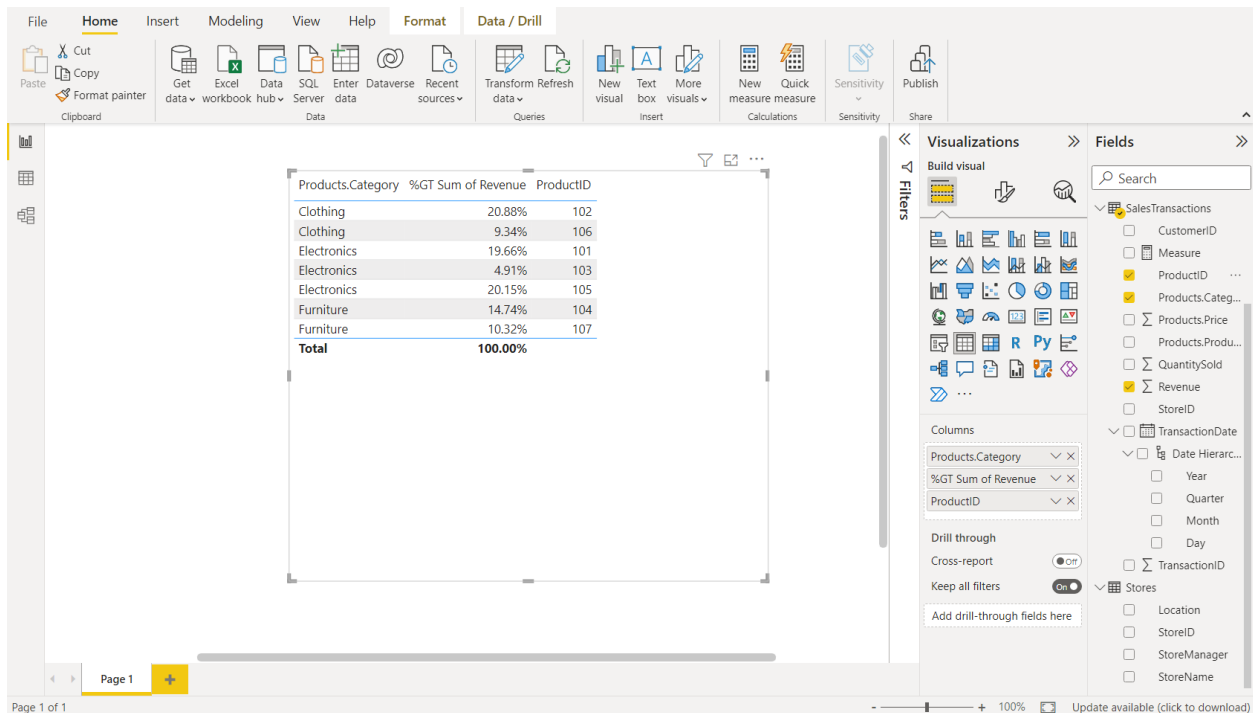
Are there specific days of the week when sales are higher?



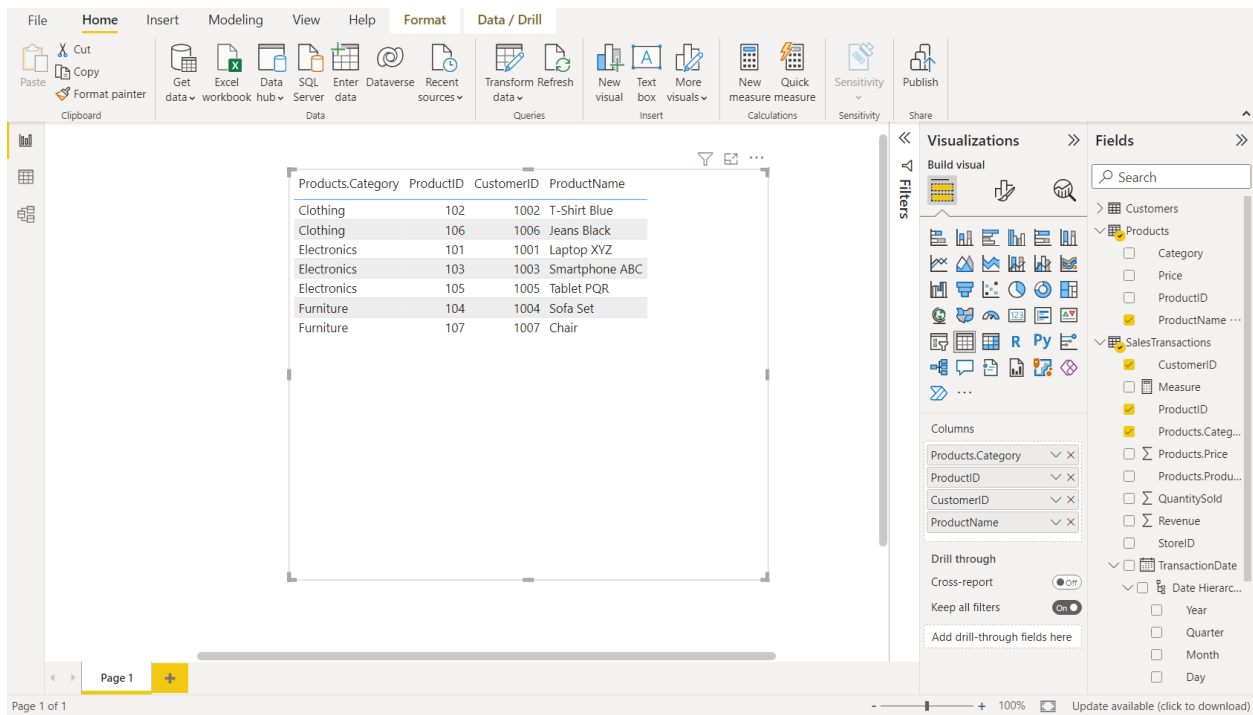
How do sales trends vary by product category on a monthly basis?



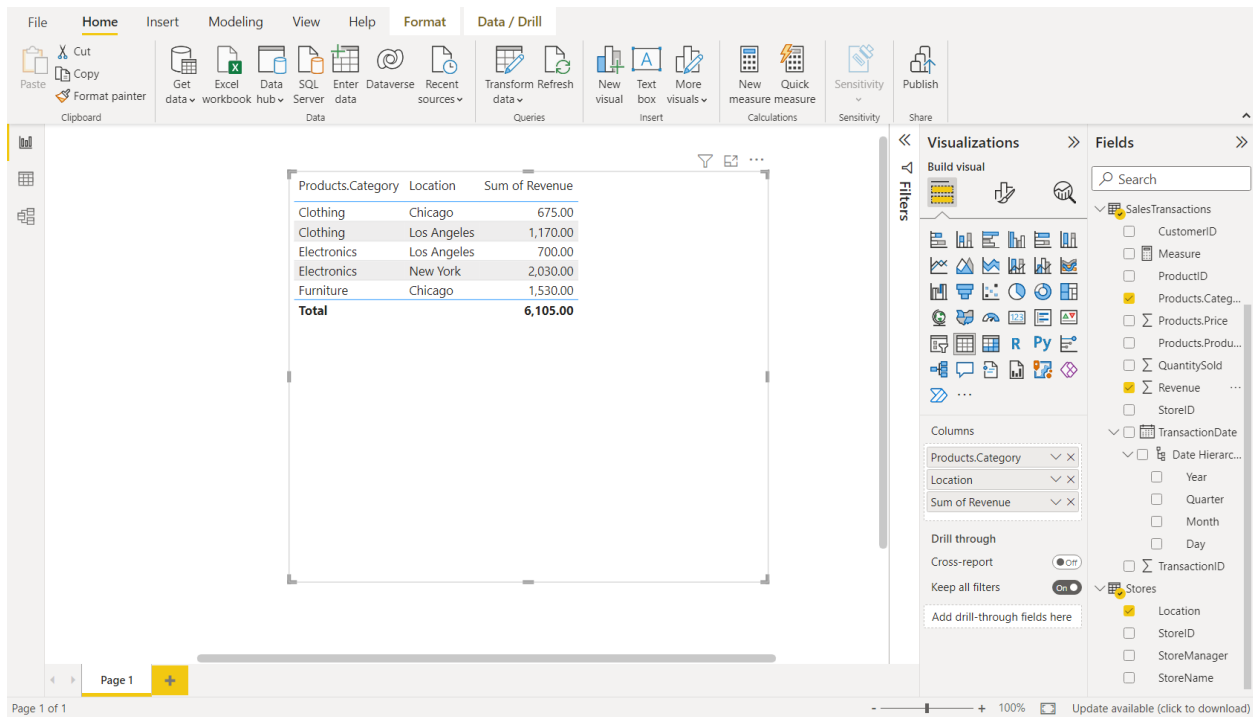
What percentage of products account for 80% of total sales revenue?



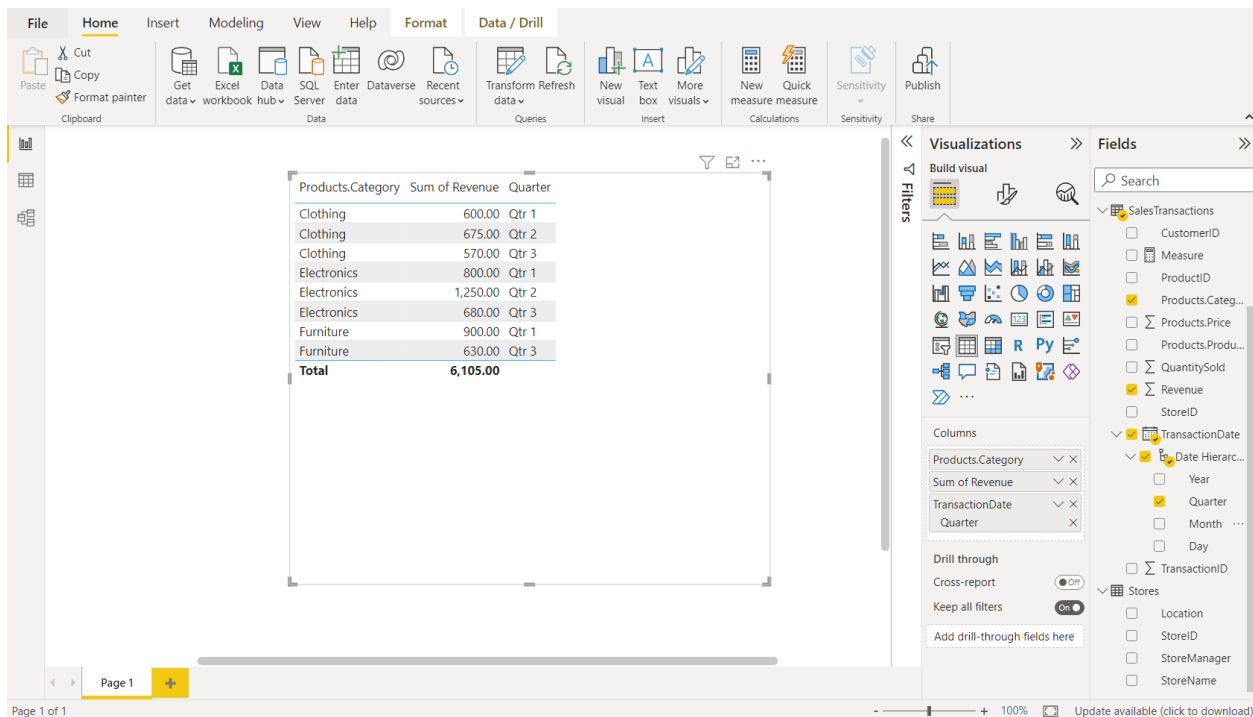
Are there any trends in repeat customer purchases?



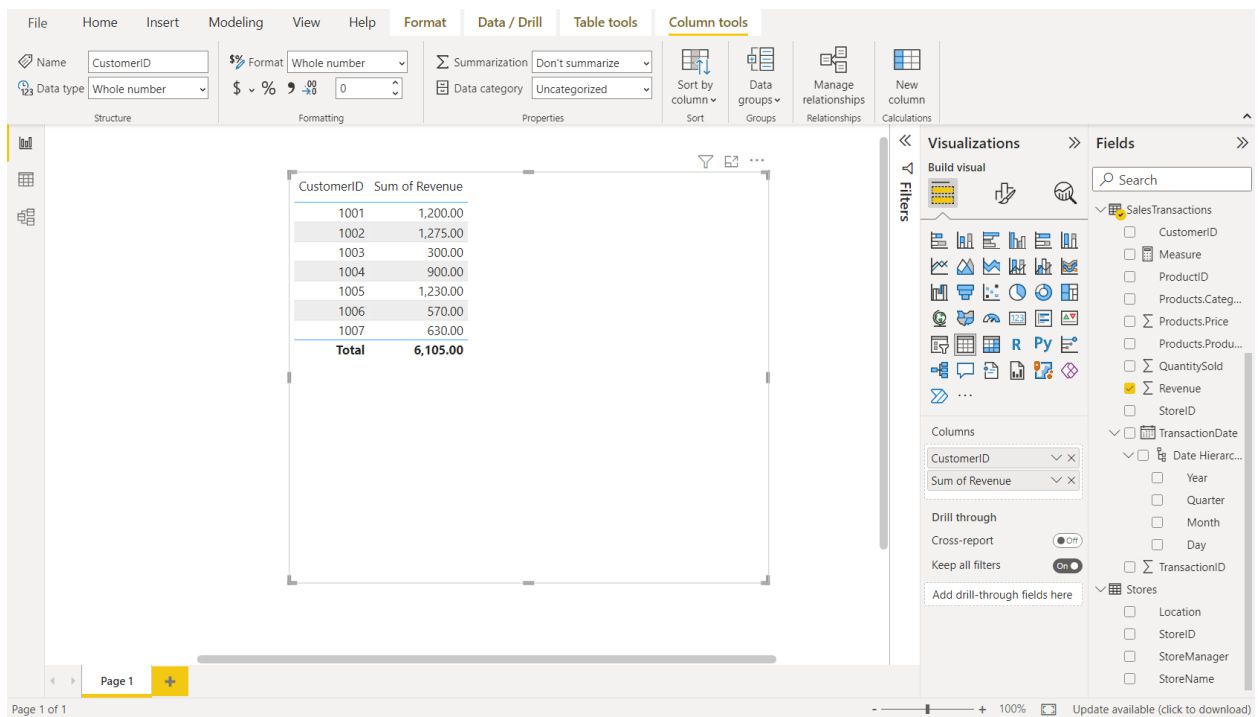
Which product categories perform best at each store location?



Are there any seasonal patterns or trends in sales for specific products or categories?

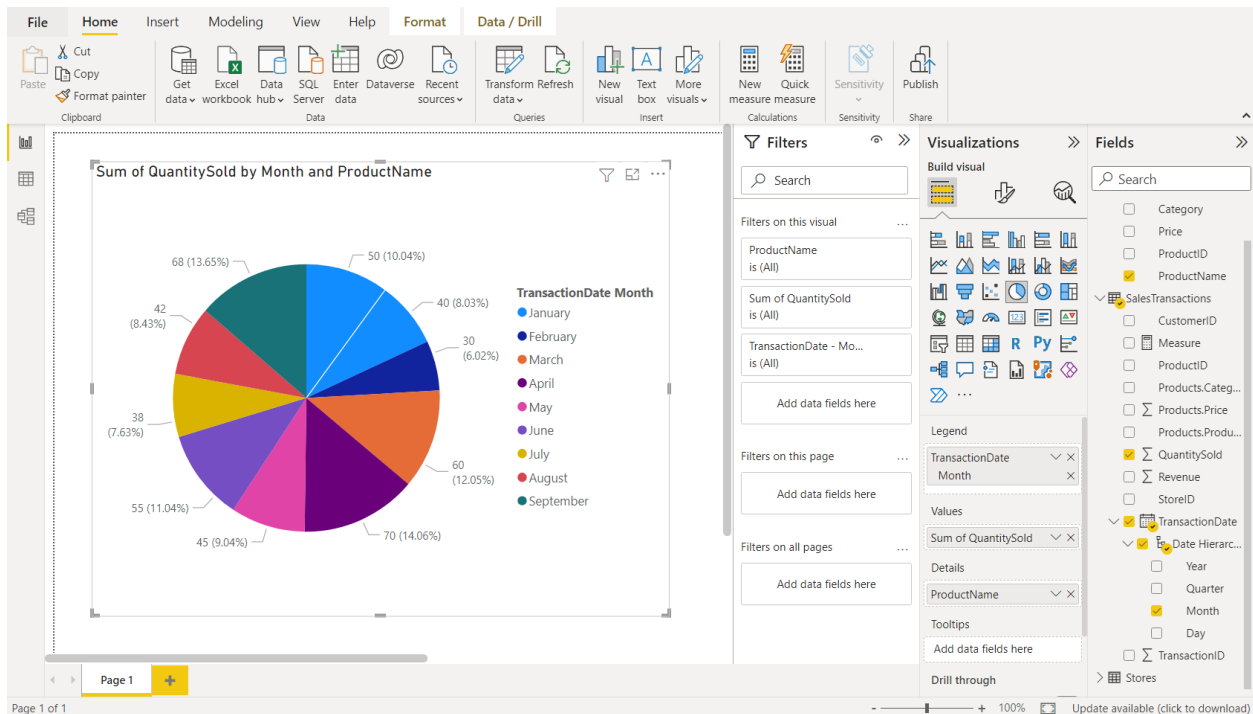


Can customers be segmented into high, medium, and low-value segments based on their purchase history

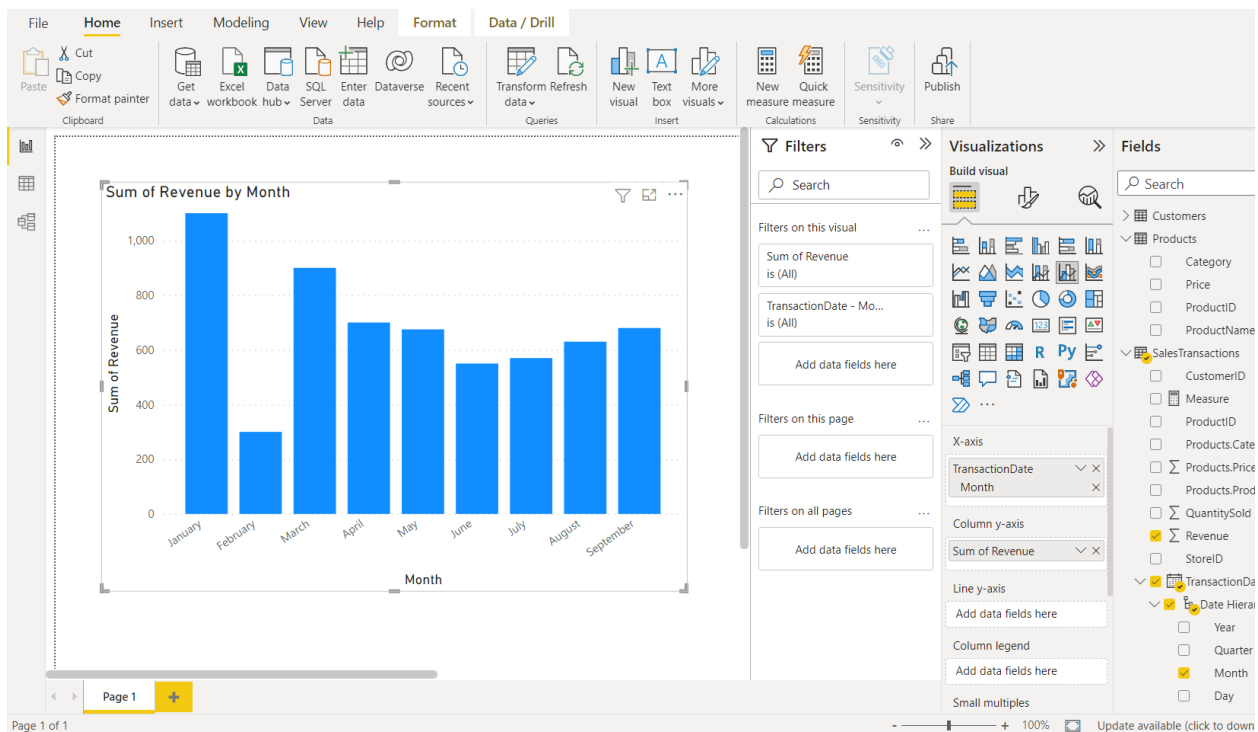


Requirement 5: Data Insights and Recommendations

Analyze Patterns: Identify patterns and trends in the data, such as seasonality or regional variations and show it or mark it.



Generate Insights: Provide actionable insights based on your analysis. For example, suggest increasing marketing efforts for the most profitable product category



Create Visual Stories: Use storytelling techniques to communicate your findings effectively in the reports and dashboards.

