

The first step for this demonstration is to use Excel Power Query to actually download the data. We found business intelligence sample data regarding products, customers, and transactions. Data is downloaded from <https://www.kaggle.com/datasets/ahmedaliraja/e-commerece-sales-data-2023-24>

### STEP I: Data Preparation –

I started by cleaning data both before and during the transformation in Power Query in the following tables:

#### Product Data

- Imported product dataset using Power Query
- Removed columns not needed for analysis
- Cleaned inconsistent pricing values (e.g., "35 on sale for 30")
- Corrected malformed price entries (e.g., "\$ 59 99")
- Removed header row incorrectly imported as data
- Converted price column to numeric/currency format
- Renamed columns for clarity and consistency
- Loaded cleaned dataset into Excel Customer Data
- Imported customer dataset using Power Query
- Removed unnecessary column(s)
- Close and load cleaned data into Excel

#### Purchase Data

- Imported dataset using Power Query
- Loaded data without transformation

#### Data Import

- Imported all datasets into a new Excel workbook using Power Query

---

### STEP II: Create Relationships

The data was in three tables: Purchases, Customers, and Transactions. They are related as follows:

<b>1. Purchases → Products</b> <ul style="list-style-type: none"><li>• Table: Purchases</li><li>• Column: ProductID</li><li>• Related Table: Products</li><li>• Related Column: ID</li></ul>	<b>2. Purchases → Customers</b> <ul style="list-style-type: none"><li>• Table: Purchases</li><li>• Column: CustomerID</li><li>• Related Table: Customers</li><li>• Related Column: ID</li></ul>
--	---

### STEP III: WORKING WITH THE DATA

Now we can create pivot tables and charts to do interaction analysis.

#### PIVOT TABLE 1. Interaction Analysis

- Created pivot table using Data Model relationships
- Rows: Product Category
- Columns: Interaction Type (View, like, or purchase a product)
- Values: Count of interactions

#### Visualization

- Pivot table showing interactions by category

#### Purpose

- Analyze customer engagement and conversion behavior across product categories

#### Findings

- Purchases lag behind likes, indicating interest but lower conversion rates
- Out of 275 interactions, purchases were significantly lower than likes
- Suggests opportunity to improve conversion rates

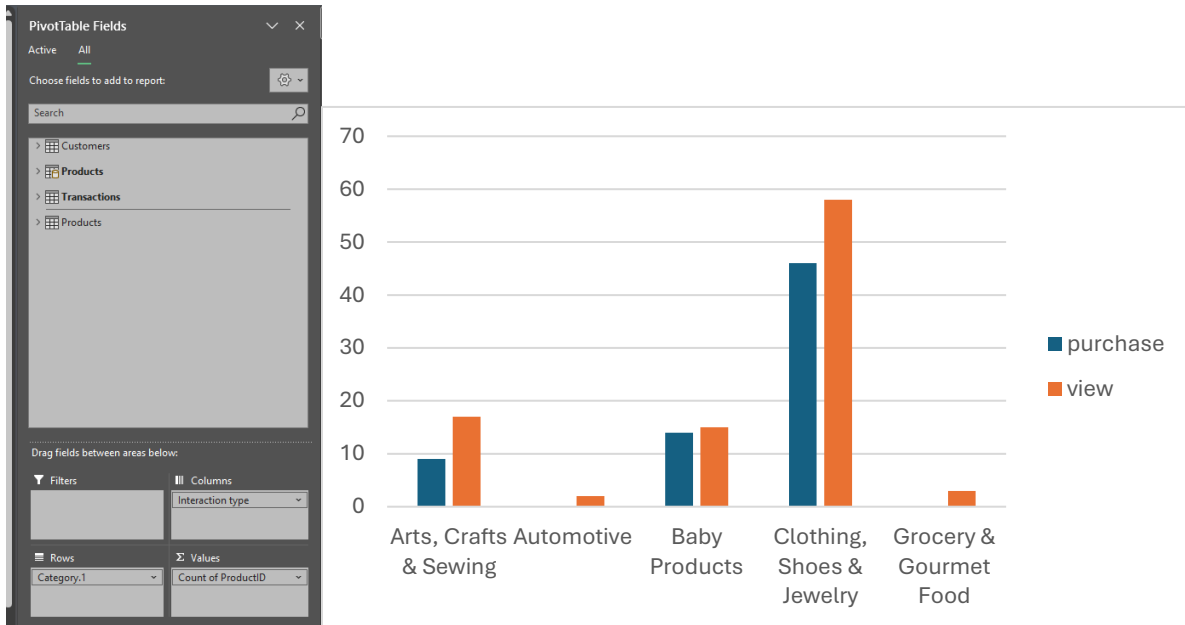
Count of ProductID	like	purchase	view	Grand Total
Arts, Crafts & Sewing   Beading & Jewelry Making   Beading Supplies   Fuse & Perler Beads	2	1	1	4
Arts, Crafts & Sewing   Beading & Jewelry Making   Beads & Bead Assortments			1	1
Arts, Crafts & Sewing   Crafting   Craft Supplies	1	1		2
Arts, Crafts & Sewing   Crafting   Craft Supplies   Pipe Cleaners			2	2
Arts, Crafts & Sewing   Crafting   Paper & Paper Crafts   Paper	1			1
Arts, Crafts & Sewing   Crafting   Paper & Paper Crafts   Paper   Construction Paper	4	3	7	14
Arts, Crafts & Sewing   Crafting   Paper & Paper Crafts   Paper   Decorative Paper	1			1
Arts, Crafts & Sewing   Crafting   Paper & Paper Crafts   Paper Craft Tools	1		1	2
Arts, Crafts & Sewing   Crafting   Weaving & Spinning   Weaving Looms			1	1
Arts, Crafts & Sewing   Crafting   Woodcrafts   Unfinished Wood		1		1
Arts, Crafts & Sewing   Painting, Drawing & Art Supplies   Drawing   Drawing Media   Crayons	2			2
Arts, Crafts & Sewing   Painting, Drawing & Art Supplies   Drawing   Drawing Media   Pencils	1	1		2
Arts, Crafts & Sewing   Painting, Drawing & Art Supplies   Painting   Kits			1	1
Arts, Crafts & Sewing   Painting, Drawing & Art Supplies   Painting   Paint Finishes		1		1
Arts, Crafts & Sewing   Painting, Drawing & Art Supplies   Painting   Paint Pens, Markers & Daubers   Pens & Markers	1			1
Arts, Crafts & Sewing   Painting, Drawing & Art Supplies   Painting   Paintbrushes   Paintbrush Sets			1	1

#### PIVOT TABLE 2: Category Analysis

The product category field contained long hierarchical text values (see the left column above), so I split category column into multiple levels using Power Query and retained first three category levels for analysis.

- **Rows:** Category Level 1 (Products table)
- **Columns:** Interaction Type (Transactions table)
- **Values:** Count of ProductID (Transactions table)

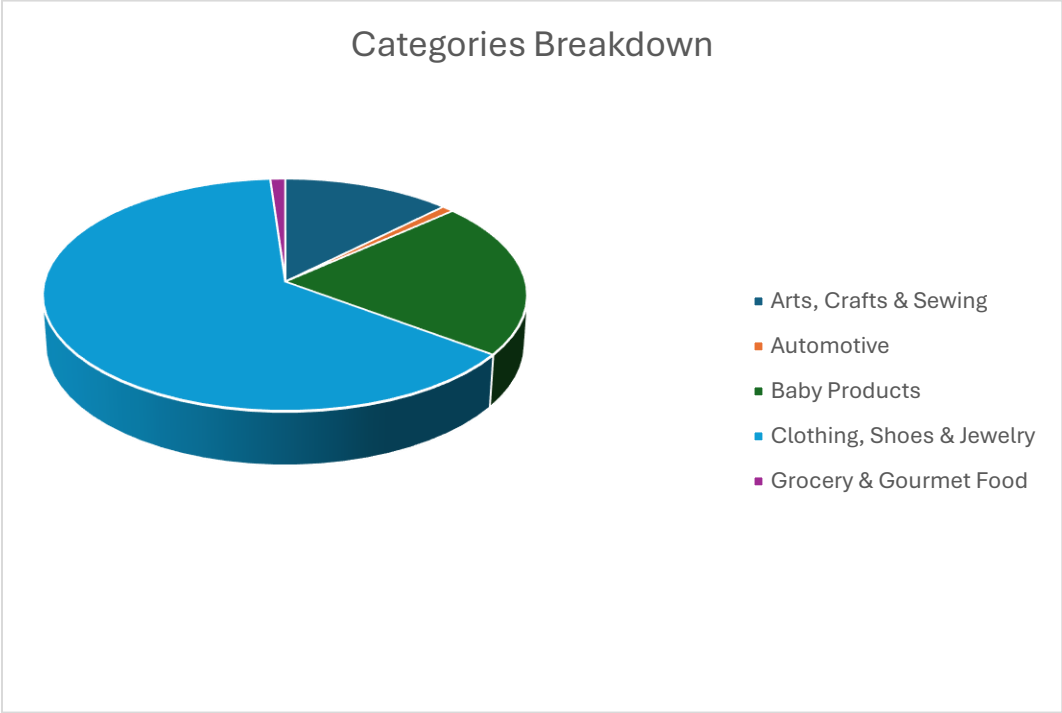
I compared Views, Likes, and Purchases across product categories, identified categories with high engagement but lower purchase conversion, and created column chart to visualize category performance



Row Labels	Count of ProductID
Arts, Crafts & Sewing	830
Automotive	124
Baby Products	9
Beauty & Personal Care	214
Cell Phones & Accessories	11
Clothing, Shoes & Jewelry	1
Electronics	630
Grocery & Gourmet Food	7
Health & Household	11
Hobbies	23
Home & Kitchen	34
Industrial & Scientific	708
Movies & TV	29
Musical Instruments	1
Office Products	2
Patio, Lawn & Garden	77
Pet Supplies	12
Remote & App Controlled Vehicle Parts	16
Remote & App Controlled Vehicles & Parts	22
Sports & Outdoors	17
Tools & Home Improvement	540
Toys & Games	17
Video Games	6662
<b>Grand Total</b>	<b>10002</b>

### Visualization — Category Distribution

I wanted to see how product categories are distributed, so I created a pie chart from the Category Distribution pivot table to visualize the proportion of interactions across product categories. This enabled quick identification of dominant product categories and made it easy to quickly highlight relative category activity levels



---

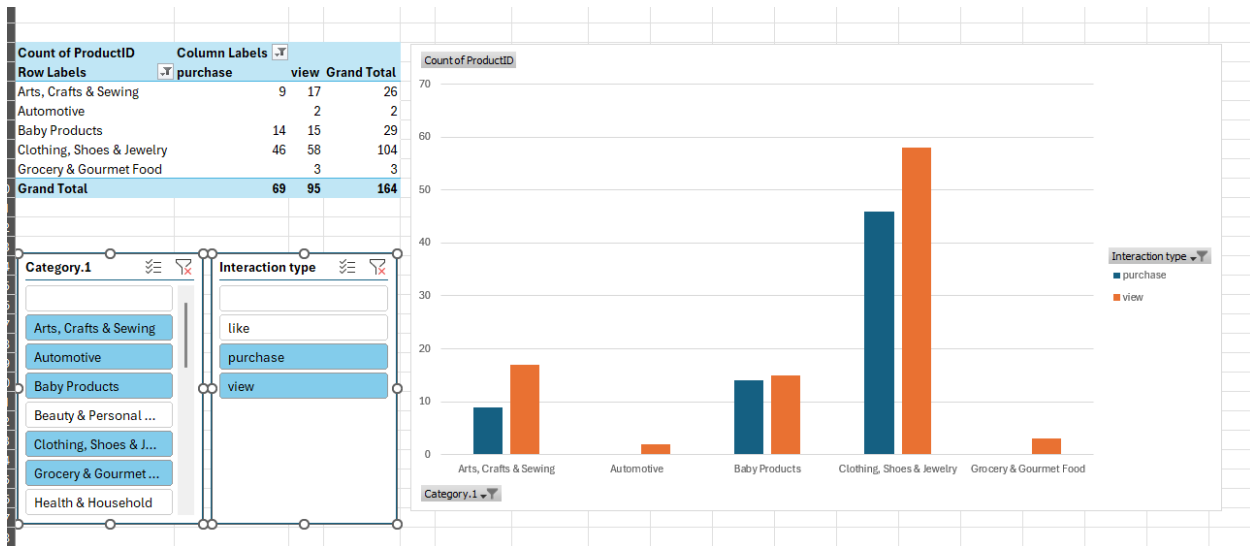
### PIVOT TABLE 3: CATEGORY HIERARCHY

Continuing the hierarchal look at categories, I took the top three categories associated with each product, enabled drill-down from broad product groupings to specific subcategories, and used hierarchical structure to identify concentration of activity within major categories. This allows us to see a hierarchical structure among categories and how they all relate. This could be a quick way to find categories that don't make sense or don't match their parent categories.

Row Labels	Count of ProductID
⊕ Arts, Crafts & Sewing	124
⊕ Automotive	9
⊖ Baby Products	214
⊖ Activity & Entertainment	10
Baby Gyms & Playmats	1
Bouncers, Jumpers & Swings	4
Walkers	4
(blank)	1
⊖ Baby Care	45
Pacifiers, Teethers & Teething Relief	45
⊖ Feeding	3
Bottle-Feeding	1
Solid Feeding	1
(blank)	1
⊖ Gifts	2
Keepsakes	2
⊖ Nursery	152
Bedding	3
Décor	133
Furniture	16
⊖ Potty Training	1
Potties & Seats	1
⊖ Strollers & Accessories	1
Accessories	1
⊖ Beauty & Personal Care	11
⊖ Foot, Hand & Nail Care	4
Nail Art & Polish	2
Tools & Accessories	2
⊖ Hair Care	2
Extensions, Wigs & Accessories	2
⊖ Makeup	2
Makeup Sets	2
⊖ Personal Care	1
Bath & Bathing Accessories	1
⊖ Skin Care	1
Lip Care	1
⊖ Tools & Accessories	1
Makeup Brushes & Tools	1
⊖ Cell Phones & Accessories	1
⊖ Accessories	1
Phone Charms	1
⊖ Clothing, Shoes & Jewelry	630

## Interactive Filtering with Slicer

Finally, I added slicer to enable interactive filtering of pivot tables, configured based on product category. This enabled dynamic filtering of hierarchical category analysis and improved usability and interactive exploration of data



## PIVOT TABLE 4: CUSTOMER LOCATION ANALYSIS

### 4. Customer Location Analysis

Finally, I created pivot table analyzing interactions by customer location.

- Rows: Customer location
- Columns: Interaction type
- Values: Count of transaction timestamps

I used this data to filter blank interaction types, applied conditional formatting color gradient to Purchase column, highlighted states with stronger and weaker purchase activity. Identified geographic differences in customer engagement

Count of Time stamp	Column Labels	like	purchase	view	Grand Total
Alabama		32.84%	37.31%	29.85%	100.00%
Alaska		38.78%	32.65%	28.57%	100.00%
Arizona		40.91%	31.82%	27.27%	100.00%
Arkansas		35.29%	41.18%	23.53%	100.00%
California		42.25%	30.99%	26.76%	100.00%
Colorado		38.18%	36.36%	25.45%	100.00%
Connecticut		34.55%	29.09%	36.36%	100.00%
Delaware		35.71%	31.43%	32.86%	100.00%
Florida		48.08%	28.85%	23.08%	100.00%
Georgia		48.08%	15.38%	36.54%	100.00%
Hawaii		36.36%	32.73%	30.91%	100.00%
Idaho		40.30%	32.84%	26.87%	100.00%
Illinois		30.77%	33.85%	35.38%	100.00%
Indiana		38.10%	34.92%	26.98%	100.00%
Iowa		48.21%	23.21%	28.57%	100.00%
Kansas		30.56%	25.00%	44.44%	100.00%
Kentucky		52.73%	29.09%	18.18%	100.00%
Louisiana		41.27%	22.22%	36.51%	100.00%
Maine		37.74%	33.96%	28.30%	100.00%
Maryland		39.71%	32.35%	27.94%	100.00%
Massachusetts		52.73%	21.82%	25.45%	100.00%
Michigan		28.30%	41.51%	30.19%	100.00%
Minnesota		42.19%	26.56%	31.25%	100.00%
Mississippi		45.61%	24.56%	29.82%	100.00%
Missouri		54.24%	25.42%	20.34%	100.00%
Montana		37.68%	31.88%	30.43%	100.00%
Nebraska		39.39%	28.79%	31.82%	100.00%
Nevada		43.75%	20.31%	35.94%	100.00%
New Hampshire		31.37%	45.10%	23.53%	100.00%
New Jersey		31.91%	34.04%	34.04%	100.00%
New Mexico		39.68%	28.57%	31.75%	100.00%
New York		37.88%	27.27%	34.85%	100.00%
North Carolina		37.29%	27.12%	35.59%	100.00%
North Dakota		35.94%	26.56%	37.50%	100.00%
Ohio		40.00%	21.67%	38.33%	100.00%
Oklahoma		39.66%	20.69%	39.66%	100.00%
Oregon		40.74%	31.48%	27.78%	100.00%
Pennsylvania		42.31%	23.08%	34.62%	100.00%
Rhode Island		33.33%	37.78%	28.89%	100.00%
South Carolina		38.98%	35.59%	25.42%	100.00%
South Dakota		35.42%	37.50%	27.08%	100.00%
Tennessee		44.44%	33.33%	22.22%	100.00%
Texas		35.29%	27.45%	37.25%	100.00%
Utah		35.29%	29.41%	35.29%	100.00%
Vermont		39.39%	27.27%	33.33%	100.00%
Virginia		35.48%	29.03%	35.48%	100.00%
Washington		47.06%	31.37%	21.57%	100.00%
West Virginia		44.44%	31.75%	23.81%	100.00%
Wisconsin		50.88%	19.30%	29.82%	100.00%
Wyoming		39.29%	32.14%	28.57%	100.00%

## CONCLUSION

This project demonstrates an end-to-end Excel-based business intelligence workflow, including data cleaning, modeling, analysis, and visualization. Using Power Query, I transformed raw e-commerce data into a structured format, then built relationships in the Data Model to enable multi-table analysis.

The analysis revealed clear patterns in customer behavior. While several product categories generate high engagement, purchases consistently lag behind views and likes, indicating opportunities to improve conversion. Hierarchical category analysis provided additional insight into how activity is distributed across product groupings, and customer location analysis highlighted geographic differences in purchasing behavior.

Interactive features such as slicers and conditional formatting were used to enhance usability and support dynamic data exploration. Overall, this project showcases the ability to use Excel as a business intelligence tool to uncover insights, communicate findings, and support data-driven decision-making.