

Retail Sales Analysis Report (Week 1)

➤ Project Objective

To analyze retail sales data from California shopping malls in order to understand **sales growth, customer behavior, revenue drivers, and churn-related patterns**, using **Excel, Power Query, SQL, and dashboarding techniques**.

➤ Week 1: Data Cleaning, Modeling & Exploratory Analysis (Excel)

✓ Data Cleaning Approach

- Removed **duplicate records and standardized column names** across datasets.
- Converted inconsistent data types:
 - Dates standardized to proper **Date format**
 - Prices, quantities, and revenue converted to **numeric values**
- Created derived columns:
 - **Revenue = Quantity × Price**
 - **Month Name & Year** extracted from invoice date
 - **Age Groups** created for demographic analysis
- Handled **missing values**:
 - Null ages grouped as “**Age Unknown**” to avoid data loss
 - Ensured no nulls in key relationship fields (customer_id, invoice_no)

✓ Data Modeling Using Power Query

- Built separate structured tables:
 - **Sales (fact table)**
 - **Customers (dimension table)**
 - **Shopping Mall (dimension table)**
- **Created relationships using *customer_id* and *shopping_mall*.**
- **Ensured one-to-many relationships by**:
 - Removing blank keys , Validating uniqueness in dimension tables

- Loaded clean, relational data into Excel Data Model for analysis.

| customer_id | gender | age | payment_method | Age_Group | sales.price |
|-------------|--------|-----|----------------|-----------|-------------|
| C150594 | Female | 60 | Cash | 46–60 | 1200.32 |
| C556552 | Female | 51 | Cash | 46–60 | 1200.32 |
| C375325 | Female | 53 | Cash | 46–60 | 1200.32 |
| C163444 | Female | 57 | Cash | 46–60 | 1200.32 |
| C247049 | Female | 54 | Cash | 46–60 | 1200.32 |
| C180375 | Female | 55 | Cash | 46–60 | 1200.32 |
| C111251 | Female | 50 | Cash | 46–60 | 1200.32 |
| C194063 | Female | 52 | Cash | 46–60 | 1200.32 |
| C101897 | Female | 60 | Cash | 46–60 | 1200.32 |
| C323574 | Female | 58 | Cash | 46–60 | 1200.32 |
| C997987 | Female | 59 | Cash | 46–60 | 1200.32 |
| C143595 | Female | 57 | Cash | 46–60 | 1200.32 |
| C265780 | Female | 59 | Cash | 46–60 | 1200.32 |
| C176271 | Female | 46 | Cash | 46–60 | 1200.32 |
| C836040 | Female | 53 | Cash | 46–60 | 1200.32 |
| C162539 | Female | 51 | Cash | 46–60 | 1200.32 |
| C260142 | Female | 57 | Cash | 46–60 | 1200.32 |
| C829141 | Female | 58 | Cash | 46–60 | 1200.32 |
| C303567 | Female | 47 | Cash | 46–60 | 1200.32 |
| C819174 | Female | 59 | Cash | 46–60 | 1200.32 |
| C279089 | Female | 57 | Cash | 46–60 | 1200.32 |
| C282163 | Female | 54 | Cash | 46–60 | 1200.32 |

| invoice_no | customer_id | category | quantity | invoice_date | price | shopping_mall | Month | Name | Year | shopping_mall.l.location | Revenue |
|------------|-------------|----------|----------|--------------|--------|------------------------|---------|------|----------|--------------------------|---------|
| I293112 | C176086 | Clothing | 2 | 1/13/2021 | 600.16 | Del Amo Fashion Center | January | 2021 | Torrance | | 1200.32 |
| I249223 | C188064 | Clothing | 2 | 1/31/2021 | 600.16 | Del Amo Fashion Center | January | 2021 | Torrance | | 1200.32 |
| I406982 | C283221 | Clothing | 2 | 1/23/2021 | 600.16 | Del Amo Fashion Center | January | 2021 | Torrance | | 1200.32 |
| I255495 | C279912 | Clothing | 2 | 1/13/2021 | 600.16 | Del Amo Fashion Center | January | 2021 | Torrance | | 1200.32 |
| I159994 | C531218 | Clothing | 2 | 1/6/2021 | 600.16 | Del Amo Fashion Center | January | 2021 | Torrance | | 1200.32 |
| I161020 | C842177 | Clothing | 2 | 1/6/2021 | 600.16 | Del Amo Fashion Center | January | 2021 | Torrance | | 1200.32 |
| I100464 | C129826 | Clothing | 2 | 1/5/2021 | 600.16 | Del Amo Fashion Center | January | 2021 | Torrance | | 1200.32 |
| I614497 | C116094 | Clothing | 2 | 1/13/2021 | 600.16 | Del Amo Fashion Center | January | 2021 | Torrance | | 1200.32 |
| I230867 | C932744 | Clothing | 2 | 1/2/2021 | 600.16 | Del Amo Fashion Center | January | 2021 | Torrance | | 1200.32 |
| I293236 | C180724 | Clothing | 2 | 1/6/2021 | 600.16 | Del Amo Fashion Center | January | 2021 | Torrance | | 1200.32 |
| I124433 | C573415 | Clothing | 2 | 1/7/2021 | 600.16 | Del Amo Fashion Center | January | 2021 | Torrance | | 1200.32 |
| I301466 | C325355 | Clothing | 2 | 1/14/2021 | 600.16 | Del Amo Fashion Center | January | 2021 | Torrance | | 1200.32 |
| I284799 | C160431 | Clothing | 2 | 1/3/2021 | 600.16 | Del Amo Fashion Center | January | 2021 | Torrance | | 1200.32 |
| I258935 | C252745 | Clothing | 2 | 1/18/2021 | 600.16 | Del Amo Fashion Center | January | 2021 | Torrance | | 1200.32 |
| I299049 | C249614 | Clothing | 2 | 1/4/2021 | 600.16 | Del Amo Fashion Center | January | 2021 | Torrance | | 1200.32 |
| I211696 | C287510 | Clothing | 2 | 1/19/2021 | 600.16 | Del Amo Fashion Center | January | 2021 | Torrance | | 1200.32 |
| I163180 | C236901 | Clothing | 2 | 1/31/2021 | 600.16 | Del Amo Fashion Center | January | 2021 | Torrance | | 1200.32 |
| I771855 | C103356 | Clothing | 2 | 1/5/2021 | 600.16 | Del Amo Fashion Center | January | 2021 | Torrance | | 1200.32 |
| I274622 | C307506 | Clothing | 2 | 1/21/2021 | 600.16 | Del Amo Fashion Center | January | 2021 | Torrance | | 1200.32 |
| I218662 | C414525 | Clothing | 2 | 1/8/2021 | 600.16 | Del Amo Fashion Center | January | 2021 | Torrance | | 1200.32 |
| I856831 | C303554 | Clothing | 2 | 1/14/2021 | 600.16 | Del Amo Fashion Center | January | 2021 | Torrance | | 1200.32 |
| I565665 | C151179 | Clothing | 2 | 1/12/2021 | 600.16 | Del Amo Fashion Center | January | 2021 | Torrance | | 1200.32 |

| shopping_mall | construction_year | area (sqm) | location | store_count |
|--------------------------|-------------------|------------|-------------|-------------|
| South Coast Plaza | 1967 | 250000 | Costa Mesa | 270 |
| Westfield Valley Fair | 1986 | 220000 | Santa Clara | 230 |
| The Grove | 2002 | 56000 | Los Angeles | 140 |
| Westfield Century City | 1964 | 133000 | Los Angeles | 200 |
| Beverly Center | 1982 | 111000 | Los Angeles | 160 |
| Fashion Valley | 1977 | 161000 | San Diego | 180 |
| Stanford Shopping Center | 1956 | 120000 | Palo Alto | 140 |
| Glendale Galleria | 1976 | 145000 | Glendale | 190 |
| Irvine Spectrum Center | 1995 | 120000 | Irvine | 130 |
| Del Amo Fashion Center | 1961 | 232000 | Torrance | 220 |

➤ Key Insights from Pivot & Dashboard Analysis

- Total Revenue: \$251.5M from 99,457 orders**
- Top Performing Mall: Del Amo Fashion Center**
- Top Revenue Category: Clothing**
- Customer Contribution:**
 - Female customers** generate ~60% of revenue
- High-Value Age Groups: 36–45 and 46–60**
- Payment Behavior: Cash** is the most used payment method



➤ Dashboard



➤ Business Findings (Combined Insights)

- Sales growth is driven mainly by **Clothing category** and **mid-aged customers**.
- Certain malls consistently **outperform others**, indicating **location-based demand strength**.
- Revenue concentration suggests **dependence on specific customer segments and malls**.
- Payment preference indicates **operational reliance on cash transactions**.

➤ Improvement Suggestions Based on Analysis

1. Customer Retention

- Introduce loyalty programs for **high-value age groups (36–60)**.

2. Category Optimization

- Expand and promote **Clothing and Technology** categories.

3. Mall-Level Strategy

- *Replicate Del Amo Fashion Center's strategies* in underperforming malls.

4. Digital Payments

- Promote **card and digital payments** to improve transaction efficiency.

5. Churn Monitoring

- *Track inactive customers* quarterly to identify churn early.

➤ Conclusion

The Excel analysis enabled effective **cleaning, structuring, and exploration** of retail sales data using **Power Query** and **Pivot Tables**. Key metrics such as revenue, orders, and customer demographics were analyzed to identify **top-performing malls and product categories**. The analysis revealed that sales are driven by **specific locations and customer segments**, with noticeable variation across **age groups and payment methods**. Overall, Excel provided clear, actionable insights and formed a strong foundation for further SQL and RFM analysis.