

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

1 customer sales shopping_mall Pivot Table Dashboard

invoice_no	customer_id	category	quantity	invoice date	price	shopping_mall	Month	Year	shopping_mall.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	C151329	Clothing	2	1/12/2021	600.16	Del Amo Fashion Center	January	2021	Torrance	1200.32

1 customer sales shopping_mall Pivot Table Dashboard

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

Total Revenue by Gender		category		Revenue by Category	
Row Labels	Sum of Revenue	Sum of Revenue	Count of invoice_no	Row Labels	Sum of quantity
Female	\$150,207,136.02		99457	Female	\$49,971,801.62
Male	\$101,298,658.23			Male	\$27,619,564.29
Grand Total	\$251,505,794.25			Grand Total	\$68,551,365.91
Revenue by Month (Trend Analysis)		Total Revenue by Shopping Mall		Revenue by Category	
Sum of Revenue	Column Labels	Row Labels	Sum of Revenue	Row Labels	Sum of Revenue
Row Labels	2021	2022	2023 Grand Total	Books	14992
January	\$9,311,207.10	\$9,917,601.79	\$6,526,787.56	Books	\$34,552.90
February	\$8,814,790.84	\$8,594,168.17	\$6,571,108.67	Clothing	103358
March	\$10,059,349.81	\$9,862,725.13	\$994,131.90	Clothing	\$113,996,791.04
April	\$9,770,141.58	\$9,816,818.11	\$1,013,615.45	Cosmetics	45465
May	\$9,767,474.25	\$9,902,187.05	\$938,616.68	Cosmetics	\$6,792,862.90
June	\$9,485,372.57	\$9,727,219.72	\$1,131,464.73	Food & Beverage	44277
July	\$10,142,596.01	\$9,862,190.26	\$916,750.45	Food & Beverage	\$949,535.05
August	\$9,490,354.67	\$9,408,352.51	\$913,654.79	Shoes	30217
September	\$8,913,202.33	\$9,803,433.46	\$731,473.04	Shoes	\$66,553,451.47
October	\$10,159,000.73	\$10,072,915.55	\$816,291.56	Souvenir	14871
November	\$9,060,094.39	\$9,154,271.83	\$485,230.74	Souvenir	\$635,814.65
December	\$9,625,306.31	\$9,914,929.50	\$559,294.01	Technology	15021
Grand Total	\$114,668,670.59	\$115,436,814.08	\$21,508,409.58	Technology	\$57,862,350.00
Payment Method Usage		Shopping Mall vs Location		Revenue by Age Group	
Row Labels	Count of invoice_no	Row Labels	Sum of Revenue	Row Labels	Sum of Revenue
Cash	44447	Costa Mesa	\$50,554,231.10	18-25	\$38,075,393.69
Credit Card	34931	Glendale	\$12,521,339.72	26-35	\$50,129,430.21
Debit Card	20079	Irvine	\$12,645,138.20	>=	\$71,734,770.01
Grand Total	99457	Los Angeles	\$62,012,808.06	Above 60	\$43,462,194.38
		Palo Alto	\$25,379,913.19	26-35	\$47,826,744.49
		San Diego	\$12,901,053.82	Age Unknown	\$277,261.47
		Santa Clara	\$24,618,827.68	Grand Total	\$251,505,794.25
		Torrance	\$50,072,481.68		
		Grand Total	\$251,505,794.25		
Top 5 customers by Revenue		Top 5 customers by Revenue		Top 5 customers by Revenue	
Row Labels	Sum of Revenue	Row Labels	Sum of Revenue	Row Labels	Sum of Revenue
C100168	\$26,250.00	C100168	\$26,250.00	C100168	\$26,250.00
C100306	\$26,250.00	C100306	\$26,250.00	C100306	\$26,250.00
C100322	\$26,250.00	C100322	\$26,250.00	C100322	\$26,250.00
C100607	\$26,250.00	C100607	\$26,250.00	C100607	\$26,250.00
C101216	\$26,250.00	C101216	\$26,250.00	C101216	\$26,250.00
Grand Total	\$131,250.00	Grand Total	\$131,250.00	Grand Total	\$131,250.00

➤ 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.