Project Objective — Retail Sales & Inventory Insights Dashboard

# **Project title:** Retail Sales & Inventory Insights (Power BI)

**Primary objective:**  
Build an interactive Power BI solution that empowers store managers and executives to monitor sales performance, track inventory health, identify best/worst-selling items, optimize reorder decisions, and forecast short-term demand. The dashboard will combine historical sales, inventory, product and store data to deliver actionable insights and KPIs that reduce stockouts, minimize overstock, and improve revenue and margin.

**Key questions the project answers**

1. Which products and categories drive the most revenue, margin, and profit across stores and channels?
2. Where and when are stockouts or overstocks occurring, and which SKUs need reordering?
3. What are sales trends (daily/weekly/monthly) and seasonality patterns by product/category/store?
4. Which customers (segments) purchase most frequently and contribute highest lifetime value?
5. How effective are promotions and discounts at driving incremental sales and profit?
6. What short-term demand forecast (next 4–12 weeks) should guide procurement and promotions?
7. Which stores or regions are underperforming and why (traffic, conversion, stock issues)?
8. What actionable recommendations (reorder points, promotion timing, markdowns) emerge from the data?

**Expected deliverables**

* Power BI report file (.pbix) with multi-page dashboard (Executive summary, Sales deep dive, Inventory health, Promotions & Margin, Forecast & Recommendations).
* Data model with relationships, calculated measures (DAX), and date/calendar table.
* Example queries / transformations (Power Query steps) and documentation of datasets and schema.
* Short recommendations document with top 5 actions to improve inventory turnover and sales.

**Success metrics / KPIs**

* Revenue, Gross Profit, Gross Margin %, Sales Growth (MoM, YoY)
* Inventory Turnover, Days Inventory Outstanding (DIO), Stockout Rate (%)
* Sell-through Rate, Fill Rate, Average Order Value (AOV)
* Promotion Lift (%) and Incremental Margin
* Forecast accuracy (MAPE) for next 4–12 weeks

# Dataset 1: Sales Transactions

**Purpose:** Core transactional record used for revenue, units sold, promotion analysis, time-series trends, customer behavior, and linking to inventory and products.

**Suggested size:** 50k–5M rows (depends on project scale; for classroom/demo 50k–200k is fine).

**Required columns (with types and short descriptions):**

1. TransactionID — *Text / GUID* — Unique identifier for each transaction (or transaction line).
2. TransactionDateTime — *DateTime* — Timestamp of the sale (useful for drill-down to hour/day/week/month).
3. StoreID — *Text / Integer* — Foreign key to Stores dataset.
4. CustomerID — *Text / Integer (nullable)* — Foreign key to Customers dataset (nullable for walk-ins).
5. SKU — *Text* — Stock-keeping unit (foreign key to Products dataset).
6. Category — *Text* — Product category (redundant field for convenience; can be derived from Products).
7. Quantity — *Whole Number* — Units sold (use negative for returns if needed).
8. UnitPrice — *Decimal / Currency* — Price per unit before discount.
9. DiscountAmount — *Decimal / Currency* — Absolute discount applied to the line (0 if none).
10. PromotionID — *Text / Integer (nullable)* — Foreign key to Promotions dataset (if sale used a promotion).
11. GrossSales — *Decimal / Currency* — Quantity \* UnitPrice (can be stored or calculated).
12. NetSales — *Decimal / Currency* — GrossSales - DiscountAmount.
13. CostPerUnit — *Decimal / Currency (nullable)* — Cost of goods sold per unit (for margin calc).
14. GrossMargin — *Decimal / Currency (nullable)* — NetSales - (Quantity \* CostPerUnit) (optional to store).
15. SalesChannel — *Text* — e.g., Online, In-Store, Phone.
16. PaymentMethod — *Text* — e.g., Cash, Card, UPI.
17. OrderType — *Text* — e.g., POS, E-commerce order, Click & Collect.
18. ReturnFlag — *Boolean / Text* — Indicates if this line is a return/refund.
19. TaxAmount — *Decimal / Currency* — Sales tax applied.
20. ShippingAmount — *Decimal / Currency (nullable)* — If applicable for online orders.
21. InvoiceNumber — *Text (nullable)* — External invoice identifier (if different from TransactionID).
22. CreatedBy — *Text (nullable)* — POS user or system that created the transaction (useful for audit).

**Notes / modelling tips:**

* Store TransactionDateTime in UTC or include timezone info; create a Date dimension table in Power BI for time intelligence.
* Prefer line-level transactions (one row per SKU per transaction). If using header + lines, you’ll need to join header and lines.
* Keep SKU and StoreID consistent with related datasets to allow proper relationships.
* For classroom/demo, include some returns (negative quantities) to show handling of refunds and netting.
* Consider privacy: if using real customers, anonymize CustomerID and PII.

# Dataset 2: Stores

**Purpose:** Provides store-level attributes for analyzing sales performance, geography, staffing, and inventory positioning. Helps with regional comparisons and performance benchmarking.

**Suggested size:** 20–200 rows (depending on the chain’s size).

**Required columns (with types and descriptions):**

1. StoreID — *Text / Integer* — Primary key (joins with Sales, Inventory).
2. StoreName — *Text* — Name of the store (e.g., “Downtown Mumbai”).
3. Region — *Text* — Larger area grouping (e.g., West, South, North).
4. City — *Text* — City where the store is located.
5. State — *Text* — State/province.
6. Country — *Text* — Country (for international comparison).
7. PostalCode — *Text* — Postal/ZIP code.
8. StoreType — *Text* — e.g., Flagship, Franchise, Outlet, Online Fulfillment.
9. OpeningDate — *Date* — Store opening date (used for cohort or age analysis).
10. ManagerName — *Text (nullable)* — Current store manager.
11. SquareFootage — *Integer (nullable)* — Store area in square feet/meters.
12. EmployeeCount — *Integer (nullable)* — Staff count (useful for sales per employee).
13. Latitude — *Decimal (nullable)* — For mapping visualizations.
14. Longitude — *Decimal (nullable)* — For mapping visualizations.
15. Status — *Text* — e.g., Active, Closed, Planned.
16. TargetSales — *Decimal / Currency* — Monthly or yearly sales target (can be used for KPI variance).
17. OperatingCost — *Decimal / Currency (nullable)* — Estimated monthly operating cost.
18. TimeZone — *Text* — Useful for aligning transaction times if multinational.
19. ChannelCoverage — *Text* — e.g., Store only, Store + Online pickup.
20. CreatedDate — *Date* — Record creation date (for audit/history).

**Notes / modelling tips:**

* Always include StoreID as a key to link with *Sales Transactions* and *Inventory*.
* Use Region and City for aggregation in Power BI maps, treemaps, and slicers.
* TargetSales is helpful for variance analysis: *Actual vs. Target*.
* Geographic fields (Latitude, Longitude) make maps more accurate than just city/state.
* For demo datasets, you can simulate ~20–50 stores with varying performance to show comparisons.

# Dataset 3: Products

**Purpose:** Master data for all products/SKUs sold, used to analyze category performance, pricing, margin, supplier dependency, and lifecycle management.

**Suggested size:** 1,000–10,000 rows (demo can use ~500–1,000).

**Required columns (with types and descriptions):**

1. SKU — *Text* — Primary key (joins with Sales, Inventory).
2. ProductName — *Text* — Product name (e.g., “Organic Basmati Rice 5kg”).
3. Category — *Text* — Broad grouping (e.g., Grocery, Electronics, Apparel).
4. SubCategory — *Text* — More granular grouping (e.g., Smartphones, Laptops, Dairy).
5. Brand — *Text* — Brand name.
6. SupplierID — *Text / Integer* — Foreign key to Suppliers dataset.
7. LaunchDate — *Date* — Date product was introduced.
8. DiscontinueDate — *Date (nullable)* — Date discontinued (if applicable).
9. UnitOfMeasure — *Text* — e.g., Each, Kg, Litre, Pack of 6.
10. StandardCost — *Decimal / Currency* — Typical procurement cost per unit.
11. ListPrice — *Decimal / Currency* — Standard selling price (before discounts).
12. MinReorderLevel — *Integer* — Minimum stock level before reorder.
13. SafetyStockLevel — *Integer* — Safety buffer to prevent stockouts.
14. LeadTimeDays — *Integer* — Average supplier lead time for replenishment.
15. GrossWeight — *Decimal (nullable)* — Weight per unit (useful for logistics).
16. Dimensions — *Text (nullable)* — Size/packaging dimensions.
17. ShelfLifeDays — *Integer (nullable)* — Useful for perishables.
18. ProductStatus — *Text* — e.g., Active, Inactive, Seasonal.
19. CreatedDate — *Date* — Record creation date (audit).
20. UpdatedDate — *Date* — Record last updated.

**Notes / modelling tips:**

* Category + SubCategory + Brand allow hierarchy drill-down in Power BI visuals.
* StandardCost and ListPrice help with Gross Margin % calculations.
* MinReorderLevel and SafetyStockLevel link naturally with the *Inventory* dataset to highlight reorder needs.
* Use ProductStatus and DiscontinueDate to filter discontinued SKUs.
* Simulated data can include both fast-moving (e.g., daily groceries) and slow-moving products (e.g., electronics).

# Dataset 4: Customers

**Purpose:** Captures customer demographics and segmentation data to analyze buying behavior, loyalty, and lifetime value. Useful for targeted marketing and sales insights.

**Suggested size:** 10,000–100,000 rows (demo ~5,000–10,000).

**Required columns (with types and descriptions):**

1. CustomerID — *Text / Integer* — Primary key (joins with Sales Transactions).
2. FirstName — *Text* — Customer’s first name.
3. LastName — *Text* — Customer’s last name.
4. Gender — *Text* — Male, Female, Other.
5. DateOfBirth — *Date* — For age calculations and cohort segmentation.
6. Email — *Text (nullable)* — Customer email (optional for demo).
7. PhoneNumber — *Text (nullable)* — Contact number.
8. Address — *Text (nullable)* — Street address.
9. City — *Text* — Customer’s city.
10. State — *Text* — Customer’s state.
11. Country — *Text* — Customer’s country.
12. PostalCode — *Text* — Postal code.
13. CustomerType — *Text* — e.g., Retail, Wholesale, Online-only.
14. JoinDate — *Date* — When customer record was created.
15. LastPurchaseDate — *Date (nullable)* — Last transaction date.
16. LoyaltyCardNumber — *Text (nullable)* — ID for loyalty programs.
17. LoyaltyTier — *Text (nullable)* — e.g., Silver, Gold, Platinum.
18. PreferredChannel — *Text (nullable)* — Online, In-Store, Both.
19. TotalSpend — *Decimal / Currency (nullable)* — Lifetime total spend (can be calculated, but useful for demo).
20. TotalOrders — *Integer (nullable)* — Total order count (again can be calculated, but speeds up demo).
21. AvgOrderValue — *Decimal / Currency (nullable)* — Derived metric.
22. ChurnFlag — *Boolean* — Indicates if customer is inactive/lost.
23. CreatedDate — *Date* — Record creation date.
24. UpdatedDate — *Date* — Last updated.

**Notes / modelling tips:**

* Combine JoinDate and LastPurchaseDate to calculate **Customer Lifetime Value (CLV)**.
* LoyaltyTier is useful for segmentation dashboards (e.g., spend by tier).
* PreferredChannel helps compare online vs in-store customers.
* For privacy, in demo datasets: anonymize Email, PhoneNumber, and Address.
* Useful DAX measures: **Active Customers (in last 90 days)**, **Churn Rate**, **CLV**.

# Dataset 5: Inventory

**Purpose:** Tracks stock levels, movements, and availability of products at each store. Helps monitor stockouts, overstocks, turnover, and reorder needs.

**Suggested size:** Daily snapshot, ~100k–1M rows (demo ~20k–50k).

**Required columns (with types and descriptions):**

1. InventoryID — *Text / Integer* — Primary key (row identifier).
2. StoreID — *Text / Integer* — Foreign key to Stores dataset.
3. SKU — *Text* — Foreign key to Products dataset.
4. InventoryDate — *Date* — Snapshot date.
5. OpeningStock — *Integer* — Units available at start of the day.
6. StockReceived — *Integer* — Units received (purchases, transfers).
7. StockSold — *Integer* — Units sold (linked indirectly to Sales Transactions).
8. StockReturned — *Integer* — Units returned (positive adjustment).
9. StockAdjusted — *Integer* — Manual adjustments (e.g., damaged, theft, audit correction).
10. ClosingStock — *Integer* — Units at end of day (Opening + Received – Sold + Returned – Adjusted).
11. OnOrderQty — *Integer* — Units currently on purchase order from supplier.
12. ReservedQty — *Integer* — Units reserved for online or bulk orders.
13. AvailableStock — *Integer* — ClosingStock – ReservedQty.
14. ReorderFlag — *Boolean* — True if AvailableStock <= MinReorderLevel (from Products).
15. StockValueCost — *Decimal / Currency* — ClosingStock \* StandardCost.
16. StockValueRetail — *Decimal / Currency* — ClosingStock \* ListPrice.
17. DaysOfSupply — *Decimal* — ClosingStock ÷ AvgDailySales.
18. StockStatus — *Text* — e.g., Healthy, Overstock, Stockout Risk.
19. CreatedDate — *Date* — Record creation date.
20. UpdatedDate — *Date* — Last update timestamp.

**Notes / modelling tips:**

* InventoryDate allows time-series tracking of stock health in Power BI.
* Use ReorderFlag + DaysOfSupply for actionable dashboards (“Reorder these SKUs”).
* StockValueCost vs StockValueRetail = valuation insights.
* ClosingStock can be validated against cumulative Sales Transactions.
* Great KPI visuals: **Stockout Rate**, **Inventory Turnover**, **Days of Supply**, **% of SKUs below reorder level**.

# Dataset 6: Suppliers

**Purpose:** Captures supplier/vendor information to analyze sourcing, lead times, costs, and supplier performance. Links to Products and Inventory for supply chain insights.

**Suggested size:** 200–2,000 rows (demo ~200–300).

**Required columns (with types and descriptions):**

1. SupplierID — *Text / Integer* — Primary key (joins with Products).
2. SupplierName — *Text* — Name of the supplier/vendor.
3. ContactName — *Text (nullable)* — Primary contact person.
4. PhoneNumber — *Text (nullable)* — Supplier contact phone.
5. Email — *Text (nullable)* — Supplier contact email.
6. Address — *Text (nullable)* — Street address.
7. City — *Text* — Supplier’s city.
8. State — *Text* — Supplier’s state/province.
9. Country — *Text* — Supplier’s country.
10. PostalCode — *Text* — Postal code.
11. LeadTimeDays — *Integer* — Average days from order to delivery.
12. MinOrderQty — *Integer* — Minimum order requirement.
13. PaymentTerms — *Text* — e.g., Net 30, Net 60, Advance.
14. SupplierRating — *Decimal (1–5)* — Performance rating (quality, delivery).
15. OnTimeDeliveryRate — *Decimal %* — % of orders delivered on time.
16. ReturnRate — *Decimal %* — % of items returned due to defects/damages.
17. PreferredSupplierFlag — *Boolean* — Indicates preferred vendors.
18. ActiveFlag — *Boolean* — True if currently active.
19. CreatedDate — *Date* — Record creation date.
20. UpdatedDate — *Date* — Last updated.

**Notes / modelling tips:**

* Use LeadTimeDays with Inventory → calculate reorder points dynamically.
* SupplierRating and OnTimeDeliveryRate enable supplier performance dashboards.
* PaymentTerms useful for finance KPIs (working capital, payables).
* PreferredSupplierFlag helps filtering in procurement dashboards.
* For demo, simulate diverse supplier performance (some high rating/low lead time, some poor rating/high lead time).

# Dataset 7: Promotions

**Purpose:** Stores information about discounts, campaigns, and special offers applied to products. Used to analyze promotion effectiveness, incremental sales, and margin impact.

**Suggested size:** 100–500 rows (demo ~100).

**Required columns (with types and descriptions):**

1. PromotionID — *Text / Integer* — Primary key (joins with Sales Transactions).
2. PromotionName — *Text* — Name of the campaign (e.g., “Diwali Mega Sale”).
3. PromotionType — *Text* — e.g., % Discount, BOGO (Buy One Get One), Bundle Offer, Coupon.
4. StartDate — *Date* — Start date of promotion.
5. EndDate — *Date* — End date of promotion.
6. ApplicableSKU — *Text (nullable)* — SKU or list of SKUs (joins with Products; can also be NULL for storewide).
7. ApplicableCategory — *Text (nullable)* — Category-level promotion.
8. DiscountPercent — *Decimal % (nullable)* — Discount % if applicable.
9. DiscountAmount — *Decimal / Currency (nullable)* — Flat discount amount.
10. BundleSKU — *Text (nullable)* — If bundle/BOGO, second SKU included.
11. Channel — *Text* — Online, In-store, Both.
12. TargetCustomerGroup — *Text (nullable)* — e.g., Loyalty Tier, New Customers, All.
13. Budget — *Decimal / Currency* — Marketing budget allocated.
14. EstimatedLiftPercent — *Decimal %* — Expected sales uplift.
15. ActualLiftPercent — *Decimal % (nullable)* — Calculated after promotion ends.
16. IncrementalSales — *Decimal / Currency (nullable)* — Sales attributed to promo.
17. IncrementalMargin — *Decimal / Currency (nullable)* — Profit from promo sales.
18. ROI — *Decimal % (nullable)* — IncrementalMargin ÷ Budget.
19. ActiveFlag — *Boolean* — True if still active.
20. CreatedDate — *Date* — Record creation date.
21. UpdatedDate — *Date* — Last updated.

**Notes / modelling tips:**

* Link promotions with **Sales Transactions** → measure promo vs non-promo sales.
* Budget vs IncrementalMargin → calculate ROI in Power BI.
* EstimatedLiftPercent vs ActualLiftPercent → measure campaign effectiveness.
* Slicers by PromotionType and Channel give quick marketing insights.
* Demo dataset can simulate festive sales (Diwali, Christmas, Black Friday) and routine promotions.

# Dataset 8: Calendar / Date Dimension

**Purpose:** Essential for all time intelligence in Power BI (YTD, MoM, rolling averages, etc.). Provides a continuous set of dates to link with Sales, Inventory, Promotions, etc.

**Suggested size:** Daily rows covering at least 5–10 years. (~3,650 rows for 10 years).

**Required columns (with types and descriptions):**

1. Date — *Date* — Primary key (joins with Sales, Inventory, Promotions).
2. Day — *Integer* — Day of month (1–31).
3. Month — *Integer* — Month number (1–12).
4. MonthName — *Text* — Full month name (January, February…).
5. MonthShort — *Text* — Abbreviation (Jan, Feb…).
6. Quarter — *Integer* — Quarter number (1–4).
7. QuarterName — *Text* — e.g., Q1, Q2.
8. Year — *Integer* — Calendar year.
9. YearMonth — *Text* — e.g., 2025-01 (for sorting charts).
10. YearQuarter — *Text* — e.g., 2025-Q1.
11. WeekOfYear — *Integer* — ISO week number.
12. DayOfWeek — *Integer* — 1=Monday … 7=Sunday.
13. DayName — *Text* — e.g., Monday.
14. DayShort — *Text* — e.g., Mon, Tue.
15. IsWeekend — *Boolean* — True if Sat/Sun.
16. IsHoliday — *Boolean* — Mark holidays (based on country).
17. HolidayName — *Text (nullable)* — Name of holiday.
18. FiscalYear — *Integer* — Fiscal year (e.g., starts in April).
19. FiscalQuarter — *Text* — e.g., FY2025-Q1.
20. IsCurrentDay — *Boolean* — Marks today.
21. IsCurrentMonth — *Boolean* — Marks current month.
22. IsCurrentYear — *Boolean* — Marks current year.

**Notes / modelling tips:**

* Generate once in Power Query or via DAX (CALENDAR or CALENDARAUTO).
* Always create relationships: Date ↔ TransactionDate, InventoryDate, PromotionDate.
* Use FiscalYear and FiscalQuarter if your project follows non-calendar fiscal cycles.
* Populate HolidayName for regional holiday sales analysis (Diwali, Christmas, etc.).