**Amazon Sales**

**Introduction**

The **Amazon Sales Report** dataset aggregates granular, time-stamped transactional data from Amazon. It tracks order details such as status, fulfillment methods, sales channel, category, size, and shipment state. This enables a comprehensive assessment of sales efficiency, product performance, and customer behavior tailored, for data-driven decision-making.

**Abstract**

This work analyzes a detailed simulation of Amazon sales data to extract actionable business insights.

* **Objective:** Evaluate sales performance across categories, channels, and fulfillment types, while examining order and shipping statuses.
* **Data & Methodology:** The dataset comprises nine key features—Order ID, Date, Status, Fulfillment, Channel, Category, Ship Service Level, Size, and Carrier Status.
* **Tools Used:** Excel for data cleaning and pivoting, SQL for relational queries and aggregations, and Power BI for interactive visualization (dashboards and reports).
* **Key Findings:** Top-performing categories are “Set” and “Kurta”; the majority of orders are shipped, with minimal cancellations or returns; fulfillment and channel-level throughput is measurable; size-wise sales visibility supports inventory planning.
* **Implications:** Insights support optimized inventory, targeted marketing, and improved logistics strategies.

**🛠️ Tools and Workflow**

| **Phase** | **Tool** | **Purpose** |
| --- | --- | --- |
| **1. Data Ingestion & Cleansing** | Excel (Power Pivot + Power Query) | Imported raw CSVs, handled missing values, standardized date formats and categories |
| **2. Data Transformation** | SQL | Used relational queries to aggregate by status, channel, category, size, and time-series trends; streamlined for Power BI import |
| **3. Analytics & Reporting** | Power BI | Created interactive dashboards to visualize KPIs (sales amount, quantity), filter by size, analyze performance across segments, and track order/shipment lifecycle |

This setup offers a scalable, repeatable framework: raw data → Excel preprocessing → SQL staging → Power BI visualization. It supports rich drilldown, performance tracking, and strategic insights

**1. Data Collection & Import**

* **Excel / Power Query**: Load raw CSV or spreadsheets. Use *Power Query* (Get & Transform) to import data into Excel
* **SQL**: Import data into a relational database (e.g., SQL Server). This supports large dataset handling and complex transformations. Many BI workflows use SQL to pre-clean and aggregate data for efficient reporting

**2. Data Cleaning & Transformation**

* **Excel / Power Query**: Clean missing values, normalize formats, duplicate using automated steps.
* **SQL**: Refine data further—write queries to join related tables (e.g. orders, categories), filter by date or status, and aggregate metrics at category, size, or status. level

**3. Data Modeling**

* **Power Pivot (Excel)**: Create internal relationships and measures (e.g., total revenue, counts) in a tabular model
* **SQL**: Optionally build summary tables or views (fact & dimension tables) to feed models downstream.

**4. Preparing Data for Reporting**

* **SQL Views**: Build optimized views with grouped metrics—e.g., revenue by category/status.
* **Power Query / Power BI**: Connect to Excel tables and SQL views. Ensure *query folding* to push transformations to the database

**5. Visualization & Dashboarding in Power BI**

* **Import / Live Connection**: Bring in data from SQL or Excel.
* **Build model**: Define relationships, hierarchies (date, category, size).
* **Create visuals**: Charts, slicers, KPIs. Pin report visuals to a dashboard.
* **Add interactivity**: Filters by size, time, and status.

**6. Testing & Validation**

* **SQL vs. DAX verification**: Verify accuracy by comparing aggregation results from SQL and DAX in Power BI .
* **Refresh checks**: Test data refresh flows—Excel→SQL→Power BI.

**7. Deployment & Sharing**

* Publish report and dashboard to Power BI Service.
* Configure scheduled refreshes for up-to-date analysis.
* Optionally, allow **“Analyze in Excel”**, enabling users to connect Excel pivot tables to the live Power BI model.

**8. Maintenance & Iteration**

* Monitor data quality and refresh logs.
* Enhance model: Add features like time intelligence, forecasting, or cohort analysis.
* Update visuals per stakeholder needs.

This workflow offers a reliable, scalable pipeline:

**Raw data → Excel (prep) → SQL (staging & aggregation) → Power BI (modeling & visualization) → Dashboard publishing & refresh.**

**✅ Conclusion**

* **Summary of Findings**: The analysis revealed that “Set” and “Kurta” categories lead in both revenue and sales volume. Most orders are fulfilled and shipped successfully, with minimal returns or cancellations. Size-wise segmentation uncovered demand concentrations useful for inventory and fulfillment planning.
* **Business Implications**: These insights empower stakeholders to:
  + Fine-tune inventory by focusing on high-demand categories and sizes.
  + Optimize fulfillment processes by monitoring shipped vs. delivered status.
  + Tailor marketing strategies to top-performing segments and channels.
* **Recommendations**:
  + **Inventory Optimization**: Prioritize stock for “Set” and “Kurta,” and use size-level visibility to maintain ideal inventory levels.
  + **Fulfillment Efficiency**: Investigate delayed or canceled shipments to minimize revenue leakage.
  + **Channel & Size-specific Marketing**: Use insights from sales channel and size trends to run targeted campaigns.
* **Next Steps**:
  + Enhance the dashboard with **forecasting features** to predict future demand.
  + Conduct **cohort or customer segmentation analysis** to understand buying behavior and improve retention.
  + Integrate **profit margin data** and customer feedback to balance sales performance with profit effectiveness.

Top of Form

Bottom of Form