Prepared by: Yogendra Raghu Varma Uppalapati

Date: June 2025

Data: 2022

Tools & Dataset Documentation

1. Dataset Overview

- Source: Amazon sales data for calendar year 2022

- Format: CSV/Excel file ("Amazon Sale Report original.xlsx")

- Rows: ~130,000 orders (Index 1–128,775)

- Columns:

- - Index: Unique order identifier
- - Date: Order date (cleaned to DD-MM-YYYY format)
- - Day/Weekday: Extracted day number and name
- - Month/Year: Extracted month and year
- - Status: Order status
- - Fulfilment: Amazon vs Merchant
- - Sales Channel: amazon.in vs other direct channels
- - Ship-Service-Level: delivery type (expedited vs standard)
- - Category: Product category (T-shirt, Shirt, Blazer, etc.)
- - Size: Product size (S, M, L, XL etc.)
- - Courier Status: Delivery outcome (Delivered, Shipped, Cancelled)
- - Qty: Quantity ordered
- - Amount: Order value (₹ INR)
- - Ship-City, Ship-State: Geography fields
- - B2B: Order type flag (Business vs Consumer)

Data Cleaning & Enhancement:

- - Removed duplicates and blank entries
- - Standardized date formats, extracted components
- - Harmonized categorical fields
- - Dropped irrelevant columns (e.g., pincode, unused metadata)
- - Engineered new fields: Weekday Name, month

2. Excel (Pivot Table Analysis)

- Tool: Microsoft Excel

- Purpose: Exploratory data analysis via pivot tables

- Actions:

- - Created 12 pivot tables covering:
- - Sales by State, City, Month, Weekday
- - Quantity by Category, Size
- - Fulfilment vs Courier Status
- - Fast vs Regular Delivery
- - B2B vs B2C split
- - Returns & Cancellations by Category

- Summarized KPIs: Total Sales, Total Orders, Average Order Value

- Outcome: Identified key trends and drill-down insights for visualization

3. Power BI (Dashboard Development)

- Tool: Microsoft Power BI Desktop

- Purpose: Interactive visual dashboard creation

- Components:

- - KPI Cards: Total Sales, Total Orders, Avg Order Value, Return Rate
- - Charts:
- - Line chart (Monthly Sales Trend)
- - Stacked bar (Product Category by Sales)
- - Filled Map (Sales by State)
- - Pie/Donut (Fulfilment Split, B2C vs B2B)
- - Bar chart (Top 20 Cities by Quantity)
- - Slicers: Category, Order Status, Delivery Type

- Styling: Amazon orange palette, consistent branding

- Interactivity: Filter-driven drill-downs

4. Word (Comprehensive Report)

- Tool: Microsoft Word (or Google Docs export PDF)

- Purpose: Detailed narrative report delivery

- Structure:

- - Title & Metadata
- - Objectives
- - Data Preparation Summary
- - Key Insights (5 subsections)
- - Strategic Recommendations
- - Conclusion
- - Appendix (Pivot list, screenshots)

- Features: Page breaks per section, headings, bullet lists, placeholders for visuals

5. Presentation (PowerPoint/Slides)

- Tool: Microsoft PowerPoint (PPTX) / Gamma or similar

- Purpose: Executive summary deck with concise, visual storytelling

- Slides:

- - Title Slide
- - Objectives
- - Data Preparation
- - Sales Performance Overview
- - Product Insights
- - Fulfilment & Delivery
- - Geographic Demand
- - Customer Segmentation
- - Recommendations
- - Conclusion
- - Appendix

- Visuals: Embedded charts from PBIX, custom infographics (data cleaning slide)

Conclusion

The above suite of tools and processes delivered a full-cycle data analysis solution—from raw data sourcing to interactive dashboard, to comprehensive report and presentation—meeting all internship deliverables.