



Amazon Sales Analysis – Documentation

1. Project Overview

This project analyzes Amazon sales data (128,976 records, 23 attributes) to uncover insights into **sales trends, customer behavior, and geographical performance**.

It also provides an **interactive Streamlit dashboard** to visualize key metrics such as revenue, quantity sold, order statuses, and geographical distribution.

2. Tech Stack Used

- **Programming Language:** Python
 - **Libraries & Frameworks:**
 - Data Processing: pandas, numpy
 - Visualization: plotly, matplotlib
 - Web App: streamlit
 - Utilities: pillow, os, glob
 - **Tools:** Jupyter Notebook, Streamlit, GitHub
-

3. Step-by-Step Setup Instructions

Clone Repository

```
git clone https://github.com/<your-username>/Amazon-Sales-Analysis.git
```

```
cd Amazon-Sales-Analysis
```

Install Dependencies

```
pip install -r requirements.txt
```

Run Jupyter Notebook

```
jupyter notebook notebooks/Amazon_Sales_Analysis.ipynb
```

Run Streamlit Dashboard

```
python -m streamlit run dashboard.py
```

4. Usage Guide

Jupyter Notebook

- Open Amazon_Sales_Analysis.ipynb
- Contains data cleaning, feature engineering, exploratory analysis, and static plots.

Streamlit Dashboard

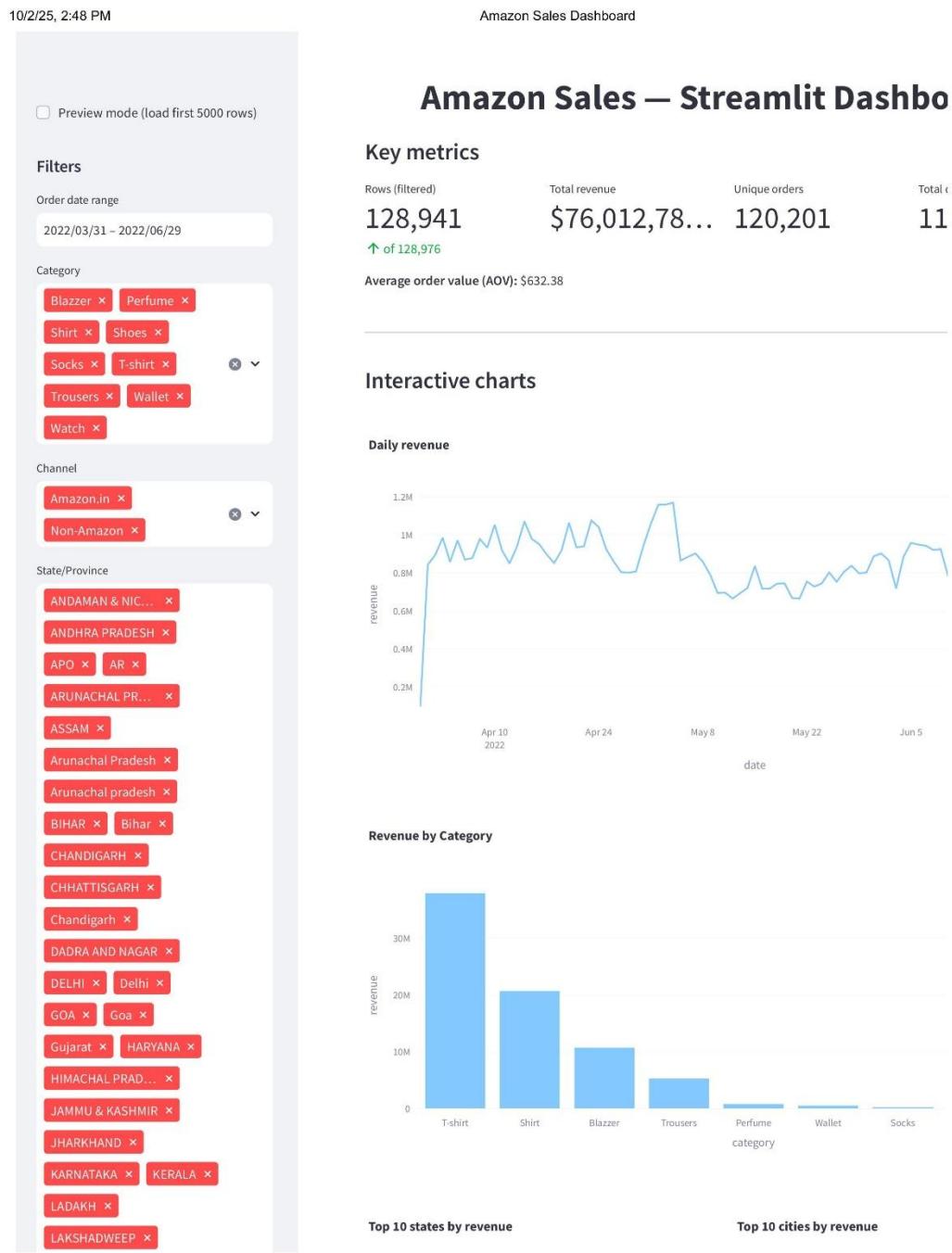
- Launch with `python -m streamlit run dashboard.py`
- Features:
 - **Filters:** Date range, category, state, city, order status, channel
 - **KPIs:** Total revenue, unique orders, total quantity sold, AOV
 - **Charts:**
 - Daily & Monthly revenue trends
 - Revenue by category, state, city
 - Order status distribution (Pie chart)
 - Revenue heatmap (state × month)
 - **Export:** Download filtered dataset as CSV

Deployment (Optional)

- Host via **Streamlit Cloud or Heroku**
 - Share the live link for direct access
-

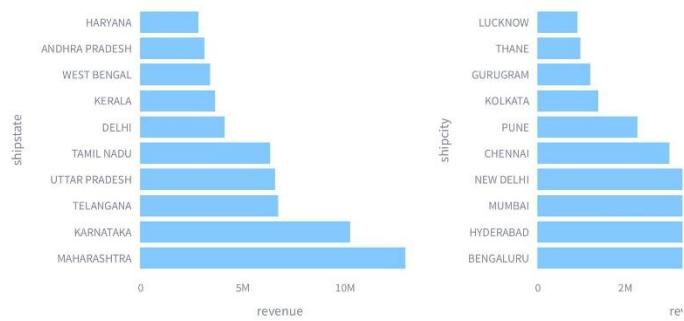
5. Key Screenshots

Dashboard Home:

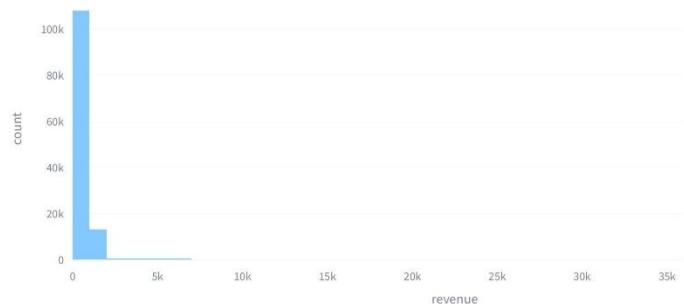


10/25/25, 2:48 PM

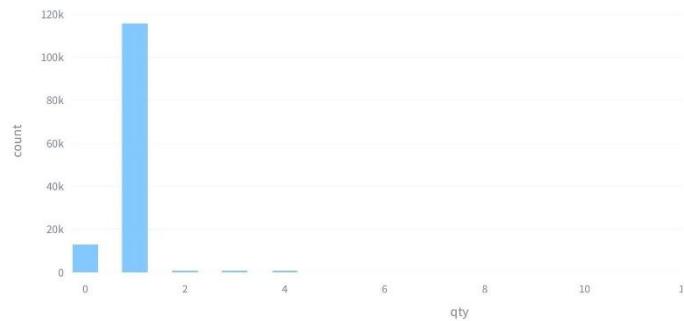
Amazon Sales Dashboard



Revenue distribution



Quantity distribution

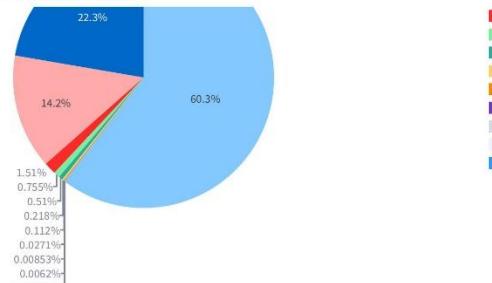


Order status share

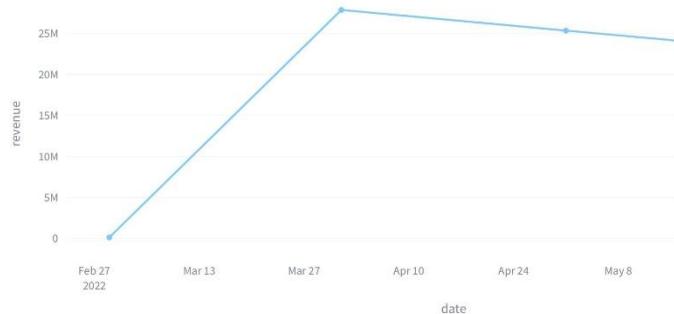


10/25, 2:48 PM

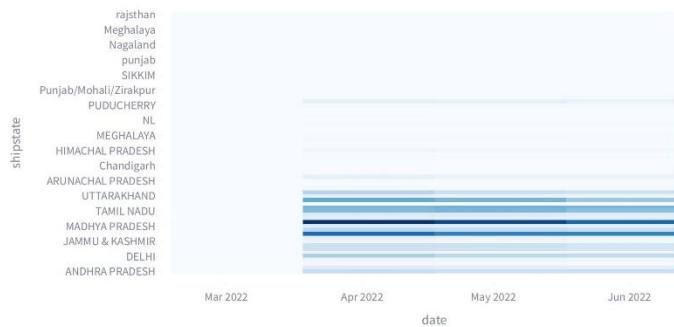
Amazon Sales Dashboard



Monthly Revenue Trend



Revenue Heatmap by State and Month



Data & export

Show filtered data table

[Download filtered CSV](#)

6. Results & Insights

- **Top Categories:** T-shirt, Shirt, Blazzer contributed >60% revenue.
 - **Top States:** Maharashtra, Karnataka, Telangana lead sales.
 - **Top Cities:** Bengaluru, Hyderabad, Mumbai drive most revenue.
 - **AOV:** ₹632.39 indicates stable purchasing patterns.
 - **Order Status:** ~14% orders cancelled → operational issue.
-

7. Additional Notes for Smooth Execution

- Ensure dataset files (amazon_sales_cleaned.csv & amazon_sales_featured.csv) are in the data/ folder.
 - If dataset is too large, use preview mode in dashboard to load limited rows.
 - Hosting the dashboard is optional but recommended.
 - For large data, running locally is smoother than free cloud deployment.
-

8. Contributors

 **Vansh Gupta**

 vanshgp2911@gmail.com

 LinkedIn :- <https://www.linkedin.com/in/vansh-gupta-3835ba26a/>