

# Data-Analytics-on-Amazon-Seller-Sales-Data

## Abstract :

To conduct analysis on our seller sales dataset and identify customer purchase interest, payment preference to generate insights and provide commercial recommendations.

## Dataset :

Sample dataset is available inside the git repository in the form of the CSV file format.

## About the data :

DATA FORMAT - Comma (',') separated text file, without quote or escape characters. The original dataset is 670 MB in size. First line in the file is header; 1 line corresponds to 1 record.

## Data Columns :

- Order Number
- Order Date
- Buyer
- Ship City
- Ship State
- SKU
- Product Description
- Product Quantity
- Item Total
- Shipping Fee
- COD (Mode of Payment online or cash)
- Delivery Status of Order

### **Tools, libraries and Languages Used:**

- Jupyter Notebook
- Python
- Pandas, Numpy, plotly, matplotlib

### **Insights:**

- Identification of the sales over dates.
- Finding most sales based on states.
- Payment method analysis for the products sold and identification of the products returned to seller.
- What was the shipping fee based on the payment method.
- Correlation Analysis.

### **Market Research :**

- Top 5 States in terms of sales are Uttar Pradesh, Maharashtra, West Bengal, Tamil Nadu, and Karnataka, Total Quantity of Orders by Top 5 States are Maharashtra, West Bengal, Tamil Nadu, Karnataka and Telangana (quantity of item yields more sales (item total)).
- Online payment is almost 3 times as big as cash on delivery payment method, Cash on delivery payment method has "Returned to seller" almost 3 times as big as online payment, We may consider promotion for online payment method.
- The fee per quantity of payment methods almost the same. (2 \$ difference is negligible, Maybe we should consider decreasing the of for shipping with online payment Or increasing fee of shipping with cash on delivery.
- Sales(item total) increases in 2021, there is an increasing trend After 2021 the sales are dropped, maybe people made too much Expenditure at the end of the year.
- The sales in December is quite high, the reason might be the Christmas, The returned to seller order status does not change much with date.
- Item total and Quantity has medium positive correlation. Shipping fee and item total has medium positive correlation. Item total vs Quantity and Item total vs Shipping Fee correlation is as expected. The number of quantity and shipping fee increases with sales.