# Amazon India: A Decade of Sales Analytics

## Skills Takeaway
Python • Pandas • Matplotlib • Seaborn • Data Cleaning • SQL • Power BI • Business Intelligence • Statistical Analysis

## Domain
E-Commerce Analytics

## Domain Introduction:
The project belongs to the E-Commerce Analytics domain, which focuses on understanding customer behavior and sales trends through data insights. It helps online retailers analyze performance metrics, customer engagement, and pricing strategies. The goal is to transform transactional data into business intelligence for strategic decision-making.

## Project Introduction:
This project analyzes 10 years of Amazon India's sales data to uncover market trends, customer patterns, and growth drivers. It provides a data-driven view of how pricing, product ratings, and regions influence long-term performance.

## Objective of the Project:
The objective is to identify key factors driving Amazon's sales performance over the decade. It aims to generate actionable insights for optimizing product strategies, pricing, and customer targeting.

## ELT Approach:
Data was Extracted from multiple CSV sources collections, Loaded into Visual Studio for structured storage, and Transformed using MySQL queries and DAX measures in Power BI for visual analytics and storytelling dashboards.

## Data Migration:

After performing data cleaning in Visual Studio, the data was migrated to MySQL by converting unstructured datasets into relational tables, enabling advanced querying and seamless integration with Power BI.

## EDA (Exploratory Data Analysis):
Analysis revealed steady growth in total sales from 2015–2024, with electronics and fashion being top-performing categories. Seasonal trends showed spikes during festive periods (Diwali, Prime Day), and South India contributed the highest regional sales.

### Feature Engineering :

Derived new variables such as average order value, customer loyalty index, and seasonal demand index to enhance trend detection and segmentation accuracy.

### Statistical Technique :

Applied a Correlation Analysis Test to explore the relationship between discount rate, rating, and sales volume — chosen to validate how price sensitivity and satisfaction affect sales outcomes.

### Conclusion:

The project concluded that discount-driven campaigns increased short-term sales but reduced profit margins, while high-rated products sustained long-term revenue growth.

### Business Suggestion / Solution:

Implement data-driven pricing models and targeted marketing for high-rating, high-conversion products. Optimize discount frequency during festive seasons to maximize profitability without over-discounting.