# 📘 Project Report

Title: Amazon E-Commerce Sales Data Analysis

Student Name: B. Navnit Kumar

Semester: 4th Semester

Department: B. Tech CSE Data Science

College: Rungta College of Engineering and Technology

Date: 18/05/2025

## 📌 1. Project Title

Amazon E-Commerce Sales Data Analysis

## 📝 2. Project Description

This project focuses on analyzing real-world Amazon sales data to extract meaningful business insights using Python. The goal is to understand customer purchasing patterns, top-selling products, trends over time, and more. It benefits online sellers, analysts, and data science learners.

## 🎯 3. Objective of the Project

- Clean and prepare raw sales data from an Amazon CSV report  
- Perform Exploratory Data Analysis (EDA) to discover trends and patterns  
- Visualize key metrics such as top categories, best-selling products, and city-wise demand  
- Integrate SQL to perform data querying and gain flexible insights  
- Generate actionable conclusions to support e-commerce decision-making

## 🔍 4. Tools & Technologies Used

- Programming Language: Python  
- Libraries: pandas, matplotlib, seaborn, sqlite3  
- Database: SQLite  
- IDE: VS Code / Jupyter Notebook  
- Platform: Kaggle for Dataset, GitHub for submission

## 📊 5. Data Analysis Steps

1. Data Loading  
 - Loaded dataset `Amazon-Sale-Report.csv` using pandas.  
2. Data Cleaning  
 - Converted date column to datetime format.  
 - Removed cancelled orders and records with missing or zero quantities.  
 - Dropped NaN values in key columns.  
 - Renamed columns for SQL and analysis.  
3. Data Exploration & Visualization  
 - Top Categories by Sales  
 - Sales Trend Over Time  
 - Top Cities by Sales  
 - Most Sold Products by Quantity  
 - Cancelled Orders Count  
4. SQL Database  
 - Inserted cleaned data into SQLite database.

## 📈 6. Visual Insights

- Bar charts for category-wise and city-wise sales  
- Line graph showing daily sales trends

## 🏁 7. Outcome

- Learned how to perform end-to-end data analysis using Python  
- Improved skills in data cleaning, visualization, and database integration  
- Discovered key business insights like top-performing categories and locations  
- Created reusable Python code and visualizations for future e-commerce projects

## 📁 8. Files Included

- Amazon\_sales\_data\_analysis.ipynb: Python script for jupyter notebook  
- Amazon-Sale-Report.csv: Raw dataset  
- requirements.txt: Python dependencies  
- README.md: Project overview and documentation  
- Amazon\_sales\_analysis.py: Raw python script

## 🔗 9. GitHub Repository

📂 https://github.com/unreasonablecoder/data-analysis-project.git