# Retail Sales Data Analysis

# (Exploratory Data Analysis)

## Project Summary

This project focuses on performing **Exploratory Data Analysis (EDA)** on a retail sales dataset to understand sales trends, customer behavior, and regional performance. The main goal was to analyze raw transactional data and convert it into meaningful insights that can help businesses make informed decisions related to marketing, inventory planning, and regional growth.

The analysis was conducted using Python in a Jupyter Notebook environment. The dataset was first cleaned and prepared to ensure accuracy, followed by detailed analysis across multiple dimensions such as region, city, and product category. Visualizations were created to clearly present trends and comparisons, making the insights easy to interpret for both technical and non-technical audiences.

---

## Project Objectives

- To clean and preprocess retail sales data for accurate analysis
- To analyze overall revenue performance across regions and cities
- To identify top-performing product categories in different regions
- To understand customer purchasing patterns at a regional level
- To find the city generating the highest revenue for the Electronics category
- To support data-driven business decisions using analytical insights

---

## Key Results & Insights

- Identified **high-revenue regions** contributing the most to overall sales
- Discovered that **product category preferences vary by region**, indicating the need for region-specific marketing strategies
- Found the **top product category in each region** based on total revenue
- Identified the **city with the highest Electronics sales**, highlighting strong demand in specific locations
- Visual analysis revealed clear sales distribution patterns across cities and categories

---

## Tools & Technologies Used

- **Python** – core programming language for analysis
- **Pandas** – data cleaning, manipulation, and aggregation
- **NumPy** – numerical computations

- **Matplotlib** – data visualization
- **Seaborn** – enhanced statistical visualizations
- **Jupyter Notebook** – interactive analysis environment

---

## Conclusion

This project demonstrates how structured EDA can transform raw retail data into actionable business insights. The analysis highlights the importance of understanding regional and category-level performance to improve sales strategies and optimize business operations. The project showcases practical skills in data analysis, visualization, and business interpretation using Python.

---