🏍️ E-Commerce Sales Data Analysis

This project is an exploratory data analysis (EDA) of a real-world e-commerce dataset using Python. It focuses on uncovering sales patterns across countries, products, and time, while demonstrating skills in data cleaning, analysis, and visualization.

📂 Dataset

- Source: Kaggle - E-Commerce Data
- Transactions: 541,909
- Time range: December 2010
- Key features: Invoice date, product description, quantity, country, customer ID

🎯 Objectives

- Clean and prepare raw sales data for analysis
- Identify top countries and products by number of sales
- Analyze sales patterns by time of day and over time
- Visualize insights through clear and professional charts

🛠️ Tools & Technologies

Tool	Purpose
Python	Programming language
Pandas	Data loading, cleaning, manipulation
Matplotlib	Data visualization (charts)
Seaborn	Statistical plots
Replit	Cloud-based development environment

📊 Key Insights

- 🇬🇧 The **United Kingdom** accounted for the vast majority of transactions
- ⏰ Most purchases occurred between **10:00 AM and 3:00 PM**
- 🧺 The most frequent products were decorative items and homeware
- 📈 Daily trends suggest sales spikes during certain periods

Visualizations

Chart	Description
Top 10 Countries	Countries with the most number of sales
Top 10 Products	Most frequently purchased items
Sales by Hour of Day	Volume of transactions across hours
Sales Over Time (by Date)	Trend of daily sales activity

Conclusions

- Peak sales occur during mid-day hours
- Marketing can be focused on top-selling items and high-sales time slots
- Improving customer identification (many missing IDs) could enable deeper segmentation
- Data cleaning (removing returns, nulls) is essential for reliable insights

---

Future Improvements

- Calculate total revenue and average order value
- Identify and analyze product returns
- Build an interactive dashboard (Power BI, Tableau)
- Apply forecasting or customer segmentation with machine learning

---

Author

Ramon Souza
Data Science Student | IT Support Specialist | Learning Python from scratch 🚀
LinkedIn • GitHub

License

This project is for educational purposes. The original dataset is publicly available on Kaggle.