Amazon Laptop Data Analysis
Overview
This project involves scraping laptop data from Amazon, cleaning and
processing the data, and performing exploratory data analysis and feature
importance analysis using Python.

Data Scraping
The data is scraped from Amazon using Selenium. The script navigates
through Amazon search result pages, extracts laptop details including
title, price, and specifications, and saves the data into a CSV file.

Script: scrape_amazon_laptops.py
Utilizes selenium and webdriver_manager to handle web scraping.
Scrapes data for up to 1500 laptops, iterating through search result
pages.
Extracts the title, price, and specifications for each laptop.
Stores the scraped data into a CSV file named amazon_laptops.csv.
Data Processing and Analysis
The data is then loaded, cleaned, and processed to extract relevant
features such as RAM, storage, and processor type. Various visualizations
are created to understand the distribution and relationships of these
features.

Script: data_analysis.py
Loading Data: Reads the CSV file containing the scraped laptop data.
Data Cleaning:
Converts the price column from string to numeric.
Extracts RAM, storage (in GB and TB), and processor type from the
specifications column.
Combines GB and TB storage values into a single storage column.
Exploratory Data Analysis:
Visualizes the distribution of laptop prices.
Compares price distributions by processor type.
Analyzes the relationship between RAM and price.
Generates a correlation matrix for price, RAM, and storage.
Feature Importance:
Uses a regression model to determine the importance of different features.
Visualizes the feature importance using a bar plot.
Visualizations
Price Distribution: Histogram showing the distribution of laptop prices.
Price by Processor Type: Box plot comparing prices across different
processor types.
Price vs RAM: Scatter plot showing the relationship between RAM and price.
Correlation Matrix: Heatmap displaying the correlation between price, RAM,
and storage.
Feature Importance: Bar plot illustrating the importance of different
features in predicting laptop prices.
Conclusion
This project demonstrates the process of web scraping, data cleaning,
exploratory data analysis, and feature importance analysis using Python.
The insights gained from the analysis can help in understanding the
factors influencing laptop prices on Amazon.