Project documentation

Laptop Products

Hamza Lamrabet
Master Wirtschaftsinformatik
lamrabet@uni-hildesheim.de

Objective

 Identify the most suitable laptop brand by considering both price and customer reviews.



Agenda



Data Cleaning and Preparation

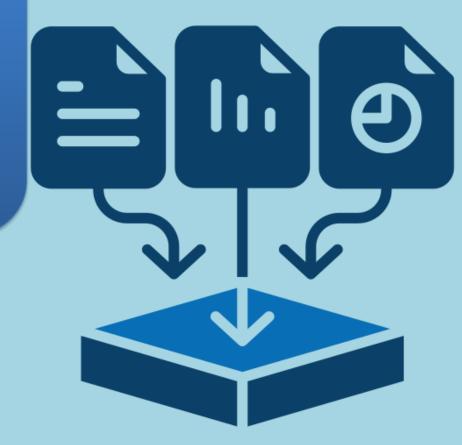
Analysis and Visualizations

Findings and Insights



Data Collection

collecting data from Amazon by scraping information about laptops, which included the laptop name, price, and customer reviews.



Tools Used

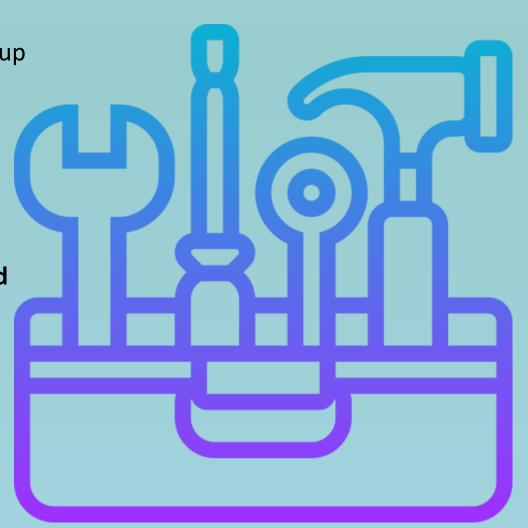
Python libraries such as BeautifulSoup for web scraping.

Pandas for data manipulation.

Requests for making HTTP requests.

CSV for saving the data.

After scraping, the data is organized and saved into a CSV file for easier analysis.



Data Cleaning and Preparation

steps in the data cleaning and preparation phase, which are essential for making the dataset ready for analysis and visualization:

- Remove Duplicates: Ensured each record is unique.
- Find and Replace Values: Corrected inconsistencies in data (e.g., "Male" vs "male").
- Categorize Data: Grouped values like income into categories (e.g., low, medium, high).

Other important cleaning tasks to focus on:

- Handling Missing Data.
- Converting Data Types.



Analysis and Visualizations

The final Power BI report contains:

 A Brand Distribution Bar Chart showing product availability by brand.

 A Rating Gauge that displays the average customer rating.

 A Price Comparison Bar Chart highlighting the average prices for each brand.

 A Summary of Total Reviews to demonstrate the strength of the data.



Findings and Insights

HP

33 products, **4,151 reviews**, **4.41 rating**, **\$1,000+** (Premium, high satisfaction).

Acer

25 products, 71,000 reviews, 4.31 rating, \$700–\$800 (Affordable, well-rated).

Lenovo

19 products, 7,067 reviews, 4.18 rating, \$800–\$900 (Balanced pricing & rating).

Dell

14 products, 79 reviews, 3.51 rating, ~\$2,000 (High-end, lower satisfaction).

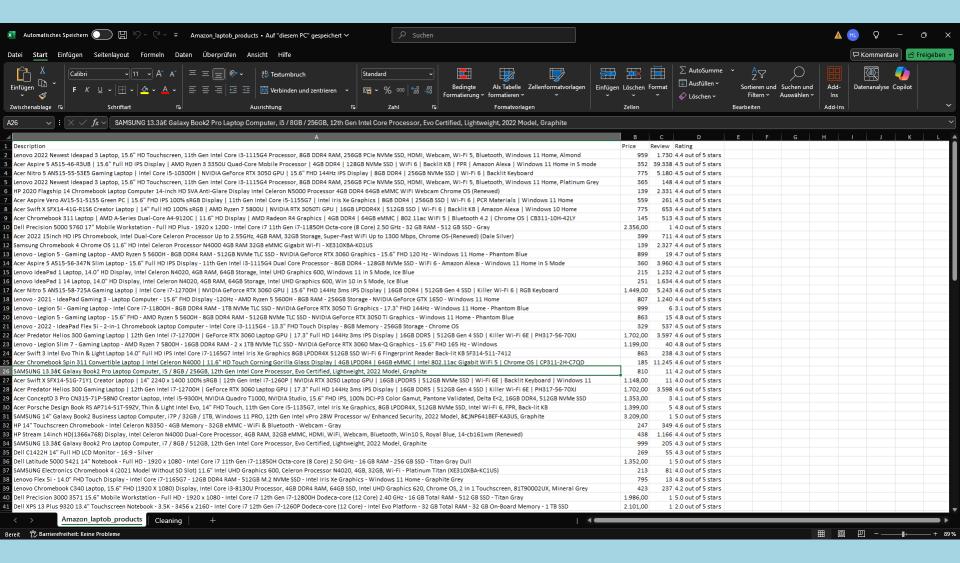
Samsung

10 products, 7,025 reviews, 4.4 rating, ~\$1,000 (Best rating, mid-range).



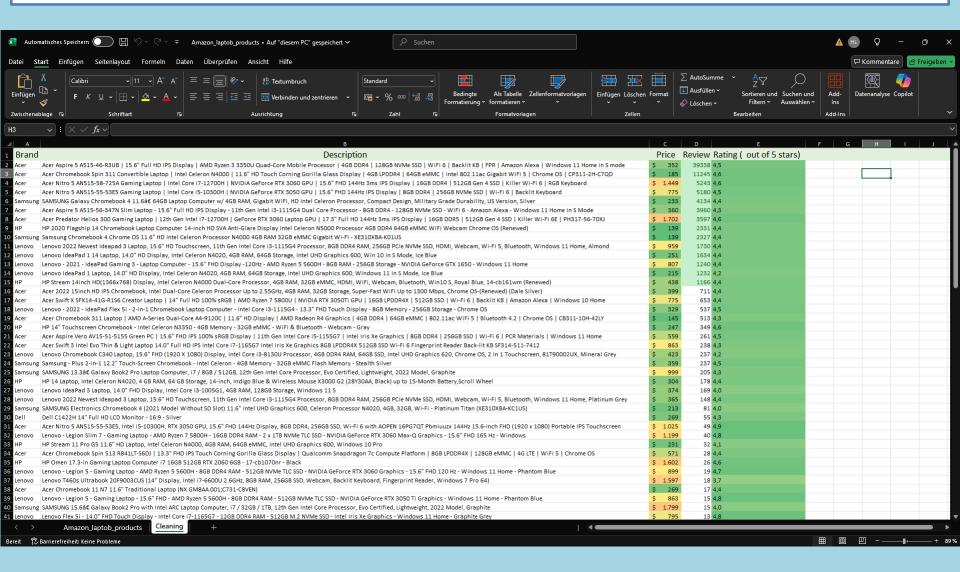
Attachment

Row Dataset



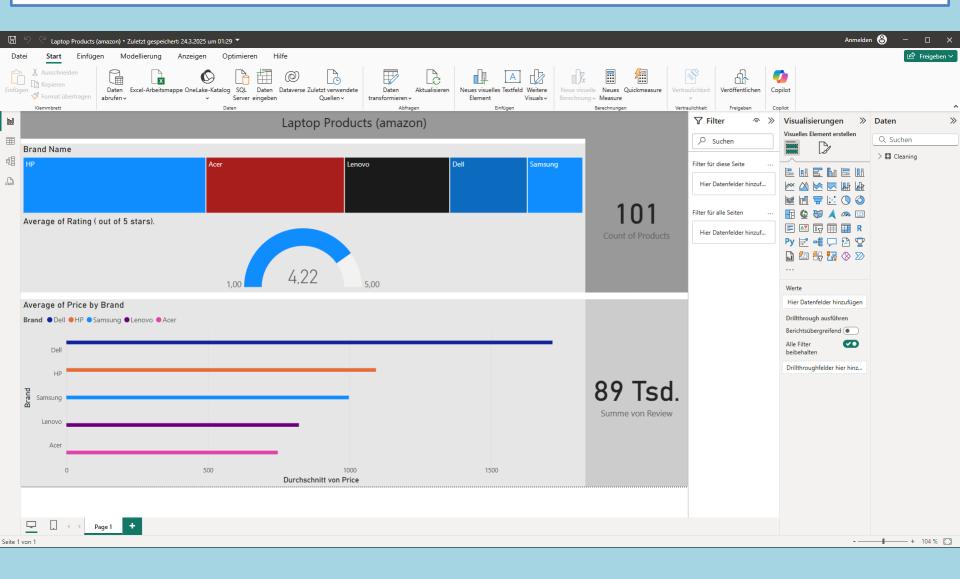
Attachment

Data Cleaning and Preparation



Attachment

Analysis and Visualizations



Thank you for your attention