LAPTOP SPECIFICATIONS DASHBOARD (March 2024)

Sweta Manandhar, Diya Shrestha, *Kathmandu University*

***Abstract*—The huge number of large datasets available in the current technological age presents unmatched possibilities for studying consumer behavior and market trends. In this advanced technological era, laptops play a vital role in every individuals life, which makes studying the dynamics of laptop prices more important for manufacturers, retailers and consumers.**

For data visualization, a potent tool, is used for transforming raw data into meaningful insights for understanding trends and patterns represented visually. This report focuses on the visualization and interpretation of the dataset, deploying an interactive dashboard built using Streamlit that contains information about various factors influencing laptop prices.

The dataset we used for the dashboard is sourced from Laptop Company Price List available on Kaggle, a data sharing platform. It contains 1303 instances and 13 attributes consisting of information like Company, Product, TypeName, RAM, Memory, OpSys, Inches, etc. The building of the dashboard employs different open-source tools and packages like Streamlit, Pandas, Xampp and Plotly.

***Index Terms*—Data Visualization, Streamlit, Pandas, Plotly**

I. INTRODUCTION

DATA Visualization is a key component of modern data analysis, providing a visual representation of complex datasets to understanding and creation of new insights. It is an essential tool for analyzing data and evaluating models at every level of a data science project. With the process of transforming unprocessed data into graphical elements like graphs, charts, and maps, visualization allows analysts to identify patterns, trends, and relationships that may not be immediately apparent from the data alone.

Before modeling, the procedure uses Exploratory Data Analysis (EDA) to examine the data and identify distribution, outliers, and fundamental statistics. For database management, we used MySQL through XAMPP. Data preparation is an important step prior to processing which involves cleaning and addressing missing values or outliers, reformatting data, making corrections to data, and combining datasets to enrich data. Depending on the data, we employ various charts, graphs, like bar charts, pie charts, scatter plots to provide a detail visualization of the dataset.

In this report, the Exploratory Data Analysis (EDA) is used to unveil significant insights from the dataset providing detailed understanding of attributes such as Product, OpSys, RAM, etc. and related factors impacting laptop prices. Similarly, Streamlit for creating the dashboard, XAMPP for dataset, Pandas for data cleaning and manipulation and Plotly for data chartings.

II. TOOLS AND PACKAGES USED

- Streamlit: Streamlit is an open-source python framework designed for creating interactive web applications for machine learning and data analysis. It is a well-liked option for quickly developing data-driven applications due to its simplicity and ease of use.

- Pandas: Pandas is a easy, open-source python library widely used for data manipulation and analysis.It allows developers to load, clean, transform and analyze datasets efficiently, making it an essential tool for data exploration and preprocessing.

- CSS (Cascading Style Sheets): CSS plays a crucial role in further enhancing the visual appeal and user experience of Streamlit applications. It allows users for fine-grained control over aspects like fonts, colors, layouts, and animations, enabling developers to create interfaces that are both visually stunning and user-friendly.

- XAMPP: XAMPP is a free and open-source cross-platform web server solution stack package developed by Apache Friends, consisting of the Apache HTTP Server and MariaDB database. It offers an easy-to-use and practical way to set up a local development environment that supports databases.

- Plotly: Plotly is an open-source python graphing package that allows us to create dynamic visuals for data visualizations like scatter plots, line charts, histogram, bar charts, etc. In addition, the interactive feature has several benefits over stationary matplotlib charts, including less time spent when analyzing the dataset.

III. DATASET SPECIFICATIONS

Kaggle is a platform for data science competitions that offers varieties of public datasets. With the advancement in technology, new laptops with different specifications are being launched quickly. To make it easier to understand and choose the best option according to one's priorities, we came up with the idea of working in laptop specifications dataset.

We extracted out dataset from https://www.kaggle.com/datasets/muhammetvarl/laptop-price. This dataset consists of 10 strings and 2 numeric values. Ten string values include: 'Company', 'Product', 'TypeName', 'ScreenResolution', 'CPU', 'RAM', 'Memory', 'GPU', 'OpSys' and

'Weight'. Two numeric values include: 'Inches' and 'price-in-rupees'. It consists of 1303 instances and 13 attributes. The dataset consists of several columns providing detailed information on different specifications of existing laptops.
The attribute 'Company' contains 19 different laptop companies that are; 'Acer', 'Lenovo', 'HP', 'Asus', 'Dell', 'Fujitsu', 'Toshiba', 'MSI', 'Apple', 'Xiaomi', 'Microsoft', 'Razer', 'Google', 'Huawei', 'Samsung', 'LG', 'Vero', 'Mediacom' and 'Chuwi'.
The attribute 'TypeName' consists of 6 different types of laptops that are; 'Notebook', 'Gaming', 'Ultrabook', '2 in 1 Convertible', 'Netbook' and 'Workstation'.
The attribute 'RAM' consists of small sizes ranging from 2GB, 4GB, 6GB, 8GB to large sizes from 12GB, 16GB, 24GB, 32GB and 64GB.
The attribute 'Inches' consists of 18 different inches including 10.1, 11.3, 11.6, 12, 12.3, 12.5, 13, 13.3, 13.5, 13.9, 14, 14.1, 15, 15.4, 15.6, 17, 17.3 and 18.4.
The attribute 'OpSys' consists of 9 different operating systems including 'Windows 10', 'No OS', 'Linux', 'Windows 7', 'macOS', 'Mac OS X', 'Windows 10 S', 'Chrome OS' and 'Android'.

A. Analysis by Filtration

In the analysis of laptop specifications dataset, there should be a proper understanding of total number of laptops of each company, average price of laptops of each company along with maximum and minimum price of laptops of each company for easier analysis. The visualization of selected companies including these attributes has been employed where the desired company's information can be carried out.

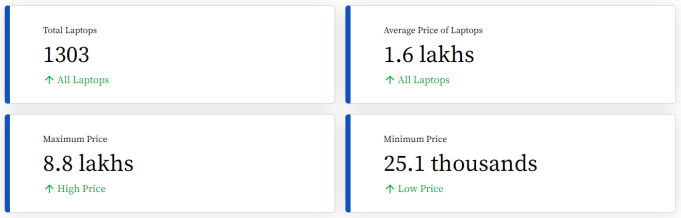

Fig. 1. Visualization of fluctuating values using filtration

B. Average Price in Rupees by Laptop Type

This visualization provides valuable insights into the pricing trends across different laptop types. It offers a comprehensive overview of the average price, measured in rupees, for various categories of laptops. By analyzing this data, users can quickly identify pricing patterns, spot outliers, and make informed decisions regarding their laptop purchases. Whether it's comparing prices between gaming laptops, ultrabooks, or workstations, this visualization serves as a crucial tool for understanding the market dynamics and optimizing budget allocations.

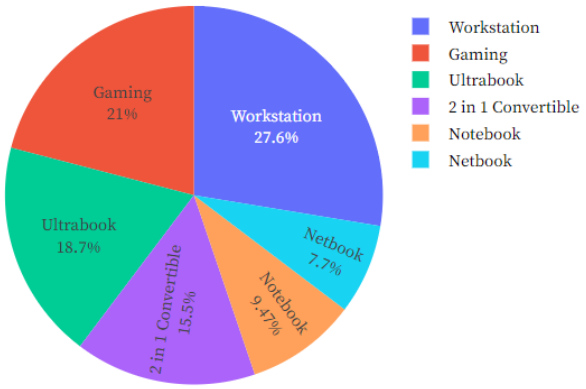

Fig. 2. Average Price in Rupees by Laptop Type using Pie-chart

C. Average Price in Rupees by RAM Capacity

This visualization provides an insightful overview of how prices vary across different RAM capacities in the Nepali market. By presenting average prices in a visual format, users can easily discern trends and fluctuations, aiding in strategic decision-making when it comes to purchasing RAM modules. This visualization not only facilitates informed choices for individual consumers looking to upgrade their systems but also offers valuable insights for businesses seeking to optimize their hardware investments.

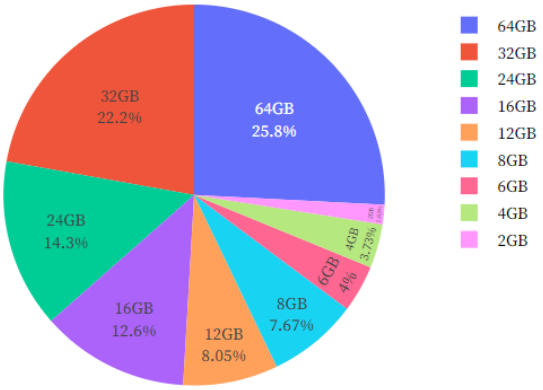

Fig. 3. Average Price in Rupees by RAM Capacity using Pie-chart

D. Price Distribution of Laptops across Companies

This visualization offers an insightful depiction of the diverse pricing landscape across various laptop brands. By presenting a comprehensive overview of price distributions, users gain valuable insights into market trends, enabling them to discern competitive pricing strategies and make informed decisions regarding their laptop purchases. This serves as a powerful tool for conducting market analysis and understanding the nuanced dynamics shaping pricing strategies.

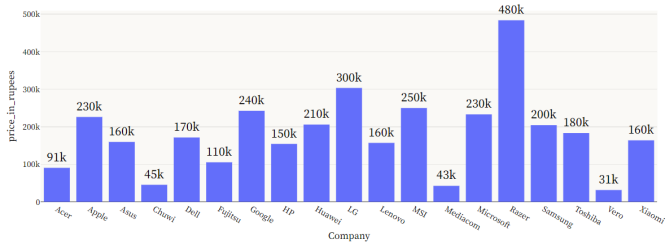

Fig. 4. Price Distribution of Laptops across Companies using Bar Chart

E. Price-wise Breakdown of Laptop Inches

This visualization provides a comprehensive overview of how laptop prices correlate with screen sizes. This visualization breaks down the cost of laptops based on the inches of their displays, enabling users to quickly discern pricing trends across different screen sizes. By visually representing this data, users can easily identify any patterns or outliers, helping them make informed decisions when selecting a laptop that meets their budget and screen size preferences.

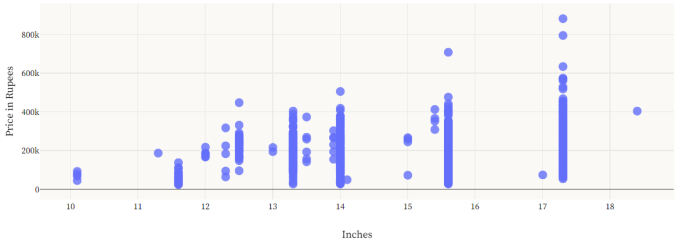

Fig. 5. Price-wise Breakdown of Laptop Inches using Scatter Plot

F. Company-wise Breakdown of Operating Systems

This visualization offers valuable insights into the distribution of operating systems such as Windows, macOS, and Linux, enabling stakeholders to identify trends, patterns, and potential areas for optimization. By presenting this breakdown by company, we gain a nuanced understanding of the technological landscape within each segment of our organization, facilitating informed decision-making and strategic planning. Whether analyzing software compatibility, security protocols, or user preferences, this visualization serves as a vital tool in enhancing operational efficiency and resource allocation across the company.

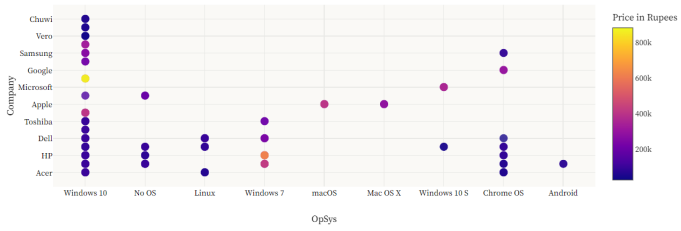

Fig. 6. Company-wise Breakdown of Operating Systems using Scatter Plot

IV. Conclusion

This report demonstrates the importance of data visualization for data-driven insights in understanding the dynamics of laptop prices. The dashboard works as a powerful tool for analytics and data visualization which makes stakeholders to make wise decisions in choosing best option as per their priorities and navigate them about the factors of competitive environment of the laptop market effectively.

Acknowledgement

First and foremost, we are immensely thankful to Assistant Professor Rajani Chulyadyo whose teaching and guidance on data mining concepts have been invaluable throughout this research. Her constructive feedback, encouragement, and willingness to engage in thoughtful discussions have significantly contributed to the refinement and completion of this research endeavor.

Furthermore, we would like to express our gratitude to the faculty and staff of the Department of Computer Science and Engineering for fostering an intellectually stimulating environment conducive to academic growth and exploration. Their support and encouragement have been invaluable throughout this journey.

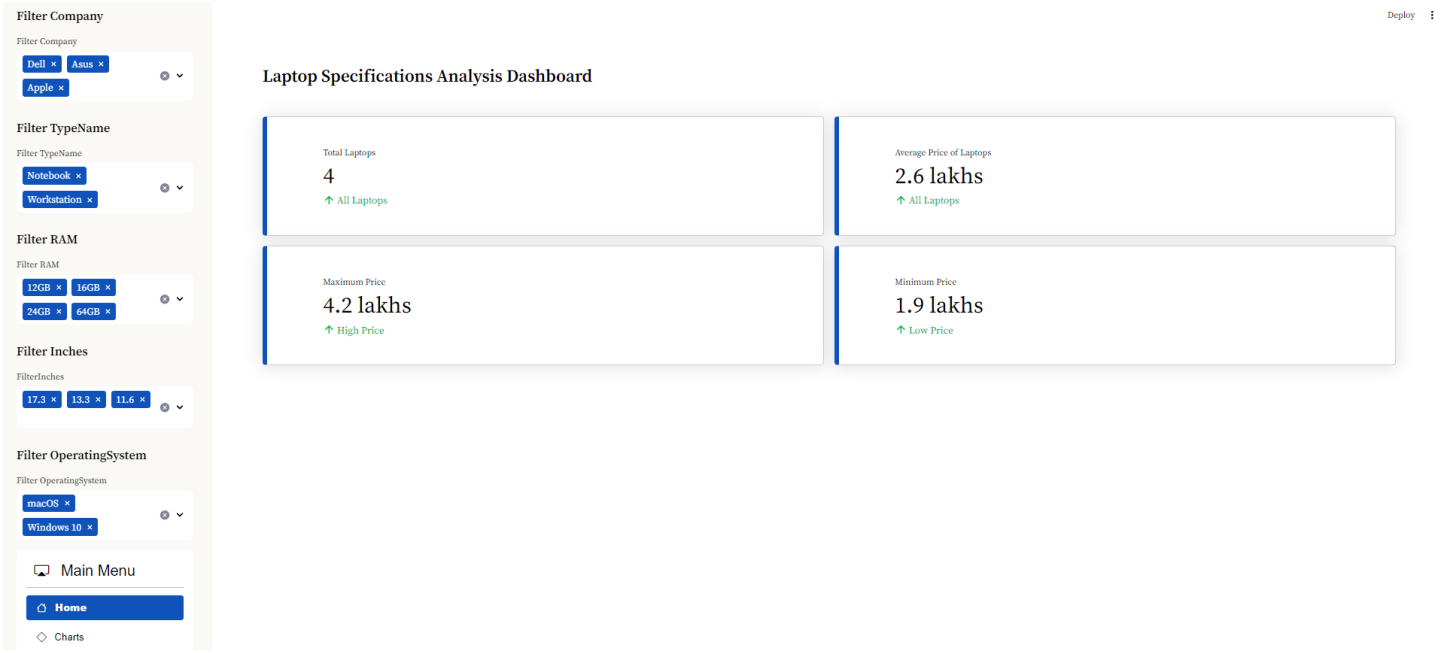

Fig. 7. First Page: Analysis by filtrating company, type name, RAM, inches and operating system

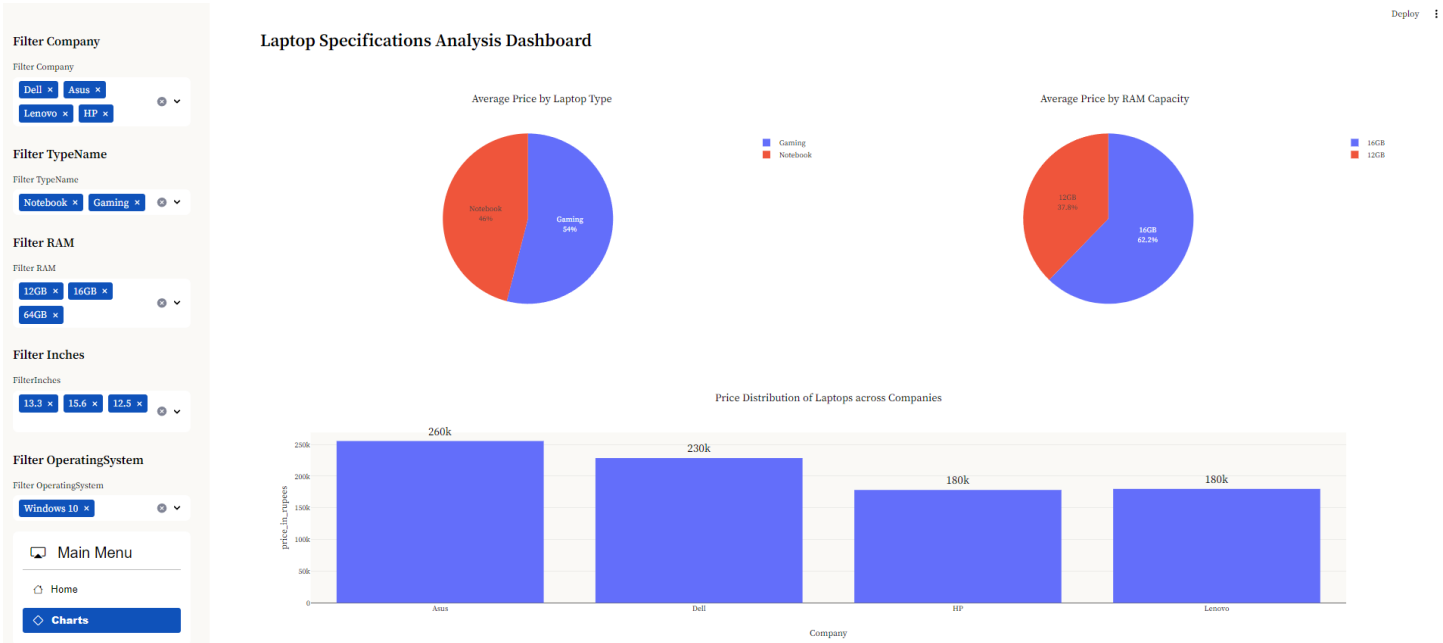

Fig. 8. Second Page: Analysis by filtration using pie-charts and bar graph

Fig. 9. Second Page: Analysis by filtration using scatter plots