

# **DATA ANALYSIS AND VISUALIZATION OF LAPTOPS**

## **A MINOR PROJECT REPORT**

*Submitted by*

**SRIKRISHNA S** **RA2111003020483**

**KALIYAN PRASANNA A** **RA2111003020486**

**GHOUSE.MD.SHAIK** **RA2111003020632**

**Under the guidance of**

**Dr. V. Gowri**

**(Assistant Professor, Department of Computer Science and Engineering)**

*in partial fulfilment for the award of the degree*

*of*

**BACHELOR OF TECHNOLOGY**

*in*

**COMPUTER SCIENCE AND ENGINEERING**

**of**

**FACULTY OF ENGINEERING AND TECHNOLOGY**



**SRM INSTITUTE OF SCIENCE AND TECHNOLOGY**

**RAMAPURAM, CHENNAI -600089**

**NOV 2024**

# **SRM INSTITUTE OF SCIENCE AND TECHNOLOGY**

**(Deemed to be University U/S 3 of UGC Act, 1956)**

## **BONAFIDE CERTIFICATE**

Certified that this project report titled “**DATA ANALYSIS AND VISUALIZATION OF BUDGET LAPTOPS**” is the bonafide work of **SRIKRISHNA S [REG NO: RA2111003020483], KALIYAN PRASANNA A [REG NO: RA2111003020486], SHAIK MOHAMMED GHOUSE [REG NO: RA211003020632]** who carried out the project work under my supervision. Certified further, that to the best of my knowledge, the work reported herein does not form any other project report or dissertation on the basis of which a degree or award was conferred on an occasion on this or any other candidate.

SIGNATURE

**Dr. V. GOWRI, M.E., Ph.D.,**

**Assistant Professor**

Computer Science and Engineering,  
SRMI Institute of Science and Technology,  
Ramapuram, Chennai.

SIGNATURE

**Dr. K. RAJA, M.E., Ph.D.,**

**Professor and Head**

Computer Science and Engineering,  
SRM Institute of Science and Technology,  
Ramapuram, Chennai.

Submitted for the project viva-voce held on \_\_\_\_\_ at

SRM Institute of Science and Technology, Ramapuram, Chennai -600089.

**INTERNAL EXAMINER**

**EXTERNAL EXAMINER**

**SRM INSTITUTE OF SCIENCE AND TECHNOLOGY**  
**RAMAPURAM, CHENNAI - 89**

**DECLARATION**

We hereby declare that the entire work contained in this project report titled “**DATA ANALYSIS AND VISUALIZATION OF BUDGET LAPTOPS**” has been carried out by **SRIKRISHNA S [REG NO: RA2111003020483], KALIYAN PRASANNA A [REG NO: RA2111003020486], SHAIK MOHAMMED GHOUSE [REG NO: RA211003020632]** at SRM Institute of Science and Technology, Ramapuram, Chennai-600089, under the guidance of **Dr. V. Gowri, Assistant Professor**, Department of Computer Science and Engineering.

**Place: Chennai**

**Date:**

**S. SRIKRISHNA**

**A. KALIYAN PRASANNA**

**SHAIK MOHAMMED GHOUSE**



**Own Work Declaration**  
**Department of Computer Science and Engineering**

**SRM Institute of Science & Technology**

**Own Work Declaration Form**

This sheet must be filled in (each box ticked to show that the condition has been met). It must be signed and dated along with your student registration number and included with all assignments you submit – work will not be marked unless this is done.

To be completed by the student for all assessments

**Degree/ Course** : B. Tech

**Student Name** : S SRI KRISHNA, SHAIK MOHAMMED GHOUSE, A KALIYAN PRASANNA

**Registration Number** : RA2111003020483, RA2111003020632, RA2111003020486

**Title of Work** : DATA ANALYSIS AND VISUALIZATION OF BUDGET LAPTOPS

I / We hereby certify that this assessment compiles with the University's Rules and Regulations relating to Academic misconduct and plagiarism, as listed in the University Website, Regulations, and the Education Committee guidelines.

I / We confirm that all the work contained in this assessment is my / our own except where indicated, and that I / We have met the following conditions:

- Clearly references / listed all sources as appropriate
- Referenced and put in inverted commas all quoted text (from books, web, etc)
- Given the sources of all pictures, data etc. that are not my own
- Not made any use of the report(s) or essay(s) of any other student(s) either past or present
- Acknowledged in appropriate places any help that I have received from others (e.g. fellow students, technicians, statisticians, external sources)
- Compiled with any other plagiarism criteria specified in the Course handbook / University website

I understand that any false claim for this work will be penalized in accordance with the University policies and regulations.

**DECLARATION:**

I am aware of and understand the University's policy on Academic misconduct and plagiarism and I certify that this assessment is my / our own work, except where indicated by referring, and that I have followed the good academic practices noted above.

If you are working in a group, please write your registration numbers and sign with the date for every student in our group.

RA2111003020483

RA2111003020632

RA2111003020486

## ACKNOWLEDGEMENT

We place on record our deep sense of gratitude to our lionized Chairman **R.SHIVAKUMAR** for providing us with the requisite infrastructure throughout the course.

We take the opportunity to extend our hearty and sincere thanks to our **Dean, Dr.M. SAKTHI GANESH, Ph.D.**, for manoeuvring us into accomplishing the project.

We take the privilege to extend our hearty and sincere gratitude to the Professor and Head of the Department, **Dr.K.RAJA, M.E, PhD.**, for his suggestions, support and encouragement towards the completion of the project with perfection.

We express our hearty and sincere thanks to our guide **Dr.V.GOWRI, M.E, Ph.D**, Computer Science and Engineering Department for her encouragement, consecutive criticism and constant guidance throughout this project work.

Our thanks to the teaching and non-teaching staff of the Computer Science and Engineering Department of SRM Institute of Science and Technology, Ramapuram Campus, for providing necessary resources for our project

## **ABSTRACT**

In order to help customers make wise purchasing decisions, this project compares all of the affordable laptops that are offered on Flipkart and Amazon, paying particular attention to important factors like functionality, cost, and user reviews. Data on key specifications laptop model, brand, processor, RAM, storage, battery life, and user feedback are gathered from each platform using web scraping tools like BeautifulSoup and Selenium. To ensure correctness and uniformity throughout the dataset, the data is then cleaned and standardized using Python's Pandas package. Using Matplotlib and Seaborn, an exploratory data analysis (EDA) reveals important insights such as pricing trends, frequent feature frequencies, and customer rating patterns.

The statistical correlations between features and user ratings are investigated using Scikit-learn, revealing information about the variables affecting customer happiness and pricing differences between the two e-commerce platforms. In order to properly convey these results, Tableau is used to build interactive dashboards that let users filter laptops based on certain criteria, such as price, performance, or battery life, which makes it simpler to find models that meet their unique needs. In addition to showcasing the best-value laptops, these dynamic visualizations let users interact directly with the data, allowing them to compare important features across platforms and find the best choices within their spending limits. This analysis ultimately provides actionable insights into budget-friendly laptop options, equipping customers with a robust tool to make well-informed purchase decisions on either platform.

# TABLE OF CONTENTS

<b>S. No.</b>	<b>TITLE</b>	<b>Page No.</b>
	<b>ACKNOWLEDGEMENT</b>	<b>v</b>
	<b>ABSTRACT</b>	<b>vi</b>
	<b>LIST OF FIGURES</b>	<b>ix</b>
	<b>LIST OF TABLES</b>	<b>x</b>
	<b>LIST OF ACRONYMS AND ABBREVIATIONS</b>	<b>xi</b>
<b>1</b>	<b>INTRODUCTION</b>	<b>1</b>
	1.1 Introduction	1
	1.2 Aim of the project	2
	1.3 Objective of the project	2
	1.4 Project Domain	3
	1.5 Scope of the Project	3
	1.6 Methodology	4
<b>2</b>	<b>LITERATURE REVIEW</b>	<b>5</b>
<b>3</b>	<b>PROJECT DESCRIPTION</b>	<b>8</b>
	3.1 Existing System	8
	3.2 Proposed System	8
	3.2.1 Advantages	8
	3.3 Feasibility Study	9
	3.3.1 Economic Feasibility	9
	3.3.2 Technical Feasibility	9
	3.3.3 Social Feasibility	9
	3.4 Software Requirement	<b>10</b>

<b>4</b>	<b>PROPOSED WORK.</b>	<b>11</b>
1.1	General Architecture	11
<b>5</b>	<b>IMPLEMENTATION AND TESTING</b>	<b>13</b>
a.	Flipkart Graphs and Dashboards	13
	Amazon Graphs and Dashboards	16
<b>6</b>	<b>RESULTS AND DISCUSSIONS</b>	<b>18</b>
a.	Efficiency of the Proposed System	19
b.	Comparison of Existing and Proposed System	20
<b>7</b>	<b>CONCLUSION AND FUTURE ENHANCEMENT</b>	<b>22</b>
a.	Conclusion	22
b.	Future Enhancements	22
<b>8</b>	<b>SOURCE CODE &amp; POSTER PRESENTATION</b>	<b>24</b>
a.	Sample Code	



## **LIST OF FIGURES**

<b>4.1</b>	<b>Architecture Diagram</b>	<b>11</b>
<b>5</b>	<b>Flipkart Graphs and Dashboards</b>	<b>13</b>
<b>5.1</b>	<b>Amazon Dashboard</b>	<b>16</b>

## **LIST OF TABLES**

<b>Table No.</b>	<b>TITLE</b>	<b>Page No.</b>
<b>1</b>	<b>Literature Survey</b>	<b>5</b>

## **LIST OF ACRONYMS AND ABBREVIATIONS:**

- 1 SQL - Structured Query Language
- 2 ETL - Extract, Transform, Load
- 3 API - Application Programming Interface
- 4 CSV - Comma-Separated Values
- 5 RAM - Random Access Memory
- 6 HDD - Hard Disk Drive
- 7 SSD - Solid State Drive
- 8 GUI - Graphical User Interface
- 9 ML - Machine Learning
- 10 EDA - Exploratory Data Analysis
- 11 BI - Business Intelligence
- 12 KPI - Key Performance Indicator
- 13 NLP - Natural Language Processing
- 14 DBMS - Database Management System

# **CHAPTER 1**

## **INTRODUCTION**

### **1.1 Introduction**

Laptops are now essential for a variety of jobs in the fast-paced digital environment we live in, including business, education, leisure, and gaming. Finding laptops that combine performance, functionality, and cost is essential for consumers given the increasing dependence on portable devices. It might be difficult to choose the best laptop within a given budget with so many options available on the market, especially for those trying to get the most for less than ₹1,00,000.

This research examines the top laptops under ₹1,000,000 that are offered in the Indian market in an effort to remedy this issue. The study guarantees the effective gathering and administration of data from various sources by utilizing SQL for data extraction, transformation, and loading (ETL) procedures. In order to enable meaningful comparisons, this methodical technique assists in gathering pertinent data, including laptop specifications, costs, and user evaluations.

In order to deliver concise, useful insights, the last step is to use Tableau to visualize the data. Customers can quickly determine which laptops provide the best performance, features, and value for their money with the help of clear dashboards and charts. Potential customers can use this research as a reference to help them choose a laptop that still fits their needs and is within their budget of ₹1,00,000.

### **1.1.1 Problem Statement**

The problem statement for this project is to address the challenges faced by consumers when selecting budget laptops from e-commerce platforms like Flipkart and Amazon. With an overwhelming variety of options, varying prices, and differing customer reviews, it becomes difficult for consumers to make informed decisions. This project aims to systematically analyze and compare budget laptops available on both platforms by leveraging data analysis and visualization techniques. The goal is to provide clear insights into pricing trends, feature availability, and customer satisfaction, thereby helping consumers identify the best value-for-money options and aiding them in making well-informed purchasing decisions.

### **1.2 Aim of the Project:**

The primary objectives of this project are to:

**Collect Data:** Gather detailed information on laptops under ₹70,000, including specifications, prices, ratings, and user reviews from leading e-commerce platforms.

**Prepare Data:** Clean and structure the collected data using SQL to ensure accuracy and consistency.

**Analyze Data:** Perform descriptive and inferential statistical analyses to identify key trends and performance metrics.

**Visualize Data:** Create interactive and insightful dashboards using Tableau to present the analysis results effectively.

**Provide Insights and Recommendations:** Identify top-performing laptops and offer recommendations tailored to various user needs based on the analyzed data.

### **1.3 Project Domain**

The domain of this project falls under e-commerce data analytics, focusing specifically on budget laptops available on major online retail platforms, Flipkart and Amazon. This domain involves extracting, analyzing, and comparing product data, including prices, specifications, customer reviews, ratings, and discounts. The goal is to help consumers and businesses make informed purchasing decisions by providing insights into the price-performance ratio, brand popularity, customer satisfaction, and platform-specific offers. E-commerce analytics combines techniques from web scraping, data mining, and data visualization to uncover trends and patterns in consumer behavior, product availability, and market competition across different platforms. This domain is critical in the current digital marketplace, where price competitiveness and customer satisfaction directly influence sales and brand reputation.

### **1.4 Scope of the Project**

The scope of this project includes gathering and analyzing data on budget laptops from Flipkart and Amazon using Python libraries for web scraping and data processing, and Tableau for visualization. It focuses on comparing key factors such as pricing, specifications, ratings, and reviews to help consumers make informed decisions. The analysis will reveal trends in customer preferences, pricing patterns, and platform-specific offers. The insights generated can be beneficial for consumers, businesses, and e-commerce platforms, marketing strategies, and competitive positioning in the budget laptop segment

## 1.5 Methodology

The first step in this project's process is gathering data using web scraping technologies like Selenium and Python's BeautifulSoup. These libraries are used to extract important data, such as laptop prices, specs, ratings, and user reviews, from e-commerce sites like Flipkart and Amazon. Finding trends in the low-cost laptop market under ₹1,000,000 is made possible by this data.

Following the collection of raw data, Pandas is used to thoroughly clean and preprocess the data. In this step, duplicates are eliminated, missing values are handled, and formats are standardized for consistent analysis. Following data cleaning, analysis is done to find important trends including consumer preferences, well-known brands, and pricing distribution. Relationships between laptop specs, such as processor type, RAM, and customer ratings, are found using statistical approaches.

Lastly, the data is visualized using Tableau through interactive dashboards and charts, allowing for product comparisons across Flipkart and Amazon. For buyers wishing to purchase a laptop for less than ₹1,00,000, these visualizations offer useful information by highlighting pricing differences, top-selling models, and user reviews. Iterative in nature, the procedure guarantees that the data remains pertinent when new items and evaluations appear.

This project's iterative data analysis and visualization approach guarantees adaptation to the ever-evolving e-commerce industry while also offering insights into low-cost computers. The data may be re-scraped and analyzed to reflect current market conditions when new laptops are introduced and customer preferences change, allowing users to stay up to date on the newest possibilities. Machine learning algorithms, including clustering and regression models, can be used to find clusters of similar laptops or forecast customer happiness based on particular features, improving the project's accuracy and relevance.

## Chapter 2

### LITERATURE REVIEW

**Table 1 : Literature Survey Data Analysis and Visualization of laptops**

S.No.	Title of the Project	Year	Name of the Journal	Inference
1	"Interactive Dashboards for Consumer Electronics"	2024	Data Visualization Journal	Showcased the effectiveness of interactive dashboards in visualizing consumer electronics data
2	"Comparative Study of Budget Laptops in 2024"	2024	Electronics Review	Compared different budget laptops based on consumer satisfaction, price, and features
3	"Trends in Laptop Sales and Specifications"	2023	Journal of Market Trends	Analysed trends in laptop sales, showing increased demand for high-performance laptops under budget constraints
4	"SQL and Tableau for Business Intelligence"	2023	International Journal of Business Analytics	Illustrated how SQL and Tableau can be integrated for effective business intelligence and data-driven decision-making



S.No .	Title of the Project	Year	Name of the Journal	Inference
5	"Comparative Study of Budget Laptops in 2024"	2024	Electronics Review	Compared different budget laptops based on consumer satisfaction, price, and features.
6	"Interactive Dashboards for Consumer Electronics"	2024	Data Visualization Journal	Showcased the effectiveness of interactive dashboards in visualizing consumer electronics data.
7	"Consumer Preferences in Laptop Purchases Post-Pandemic"	2022	Journal of Consumer Behavior	Identified shifts in consumer preferences for laptops, with increased emphasis on specifications and online reviews
8	"Impact of Online Reviews on Electronics Purchases"	2022	Journal of Marketing	Found that online reviews significantly impact consumer purchasing decisions, particularly for electronics

<b>S.No.</b>	<b>Title of the Project</b>	<b>Year</b>	<b>Name of the Journal</b>	<b>Inference</b>
9	"Interactive Dashboards for Consumer Electronics"	2024	Data Visualization Journal	Showcased the effectiveness of interactive dashboards in visualizing consumer electronics data.
10	“Data analytics and visualization using Tableau utilitarian for COVID-19 (Coronavirus)” (Base Paper)	2020	Global Journal of Engineering and Technology Advances	Analyzed trends in Covid Data  Showing decline in people getting affected after finding vaccine for Coronavirus
11	“Comparative Analysis of E-commerce Platforms Using Consumer Review Data”	2020	Journal of Retailing and Consumer Services	Features like price, reliability, and customer service were key factors in positive customer ratings across various product types.
12	“Market Analysis of Budget Laptops through Feature Preferences and Pricing Trends”	2019	International Journal of Information Management	Found that features like RAM, battery life, and screen resolution significantly influence purchasing decisions in budget segments.

## **Chapter 3**

### **PROJECT DESCRIPTION**

#### **3.1 Existing System**

The existing system for comparing budget laptops on e-commerce platforms like Flipkart and Amazon is largely manual, where users have to individually browse through numerous products, compare specifications, prices, and read reviews on each platform. There is no centralized platform or tool that automates this comparison or provides detailed insights into pricing trends, customer satisfaction, or performance across different brands. The lack of automated analysis makes decision-making time-consuming and often confusing for users.

#### **3.2 Proposed System**

The proposed system automates the process of data collection, analysis, and visualization of budget laptops from Flipkart and Amazon. Using Python libraries for web scraping, data from both platforms will be gathered, cleaned, and analyzed, offering a detailed comparison of prices, specifications, and customer reviews. Visualizations created in Tableau will display key insights such as price trends, brand popularity, and user ratings, simplifying decision-making for consumers and enabling businesses to understand market dynamics more effectively. This system aims to provide real-time, data-driven insights for both users and e-commerce stakeholder

##### **3.2.1 Advantages**

- i. Automated data collection and analysis from multiple platforms.
- ii. Enhanced informed decision-making for consumers.
- iii. Data-driven insights into pricing trends and consumer preferences.
- iv. Streamlined comparison of laptop specifications and customer reviews.

### **3.3 Feasibility Study**

A Feasibility study is carried out to check the viability of the project and to analyze the strengths and weaknesses of the proposed system. The application of usage of mask in crowd areas must be evaluated. The feasibility study is carried out in three forms

- a) Economic Feasibility
- b) Technical Feasibility
- c) Social Feasibility

#### **3.3.1 Economic Feasibility**

The proposed system does not require any highcost equipment. This aspect considers the cost-effectiveness of the project. The main expenses include setting up web scraping tools, software for data analysis (like Python libraries), and Tableau for visualization. Since these tools are either open-source or relatively low-cost compared to their benefits, the project is economically viable. Additionally, the time and effort saved by automating the comparison process for users justifies the investment.

#### **3.3.2 Technical Feasibility**

The technical feasibility examines whether the technology needed for the project is available and accessible. Python libraries like BeautifulSoup, Selenium for web scraping, Pandas for data analysis, and Tableau for visualization are widely available and well-documented. The project requires basic-to-intermediate programming and data analysis skills, which are manageable with existing resources. Hence, the project is technically feasible.

#### **3.3.3 Social Feasibility**

This evaluates the acceptance and impact of the system on users and society. Consumers often face challenges comparing laptops across multiple e-commerce platforms, and this project offers a solution by providing detailed, data-driven insights. The project is likely to be well-received as it improves user decision-making, reduces

confusion, and enhances the shopping experience. Thus, it is socially feasible and beneficial to the target audience.

### **3.4 System Specification:**

(i) Basic Normal Laptop

(ii) Good Stable Internet Connection

#### **3.4.1 Software Specification**

##### **Data Processing and Storage:**

**Python:** Programming language for data processing, cleaning, and analysis. Libraries like Pandas and NumPy will be useful.

**SQL:** Structured Query Language for managing and querying relational databases.

**MySQL:** Popular relational database management systems for storing the collected and processed data.

##### **Data Analysis:**

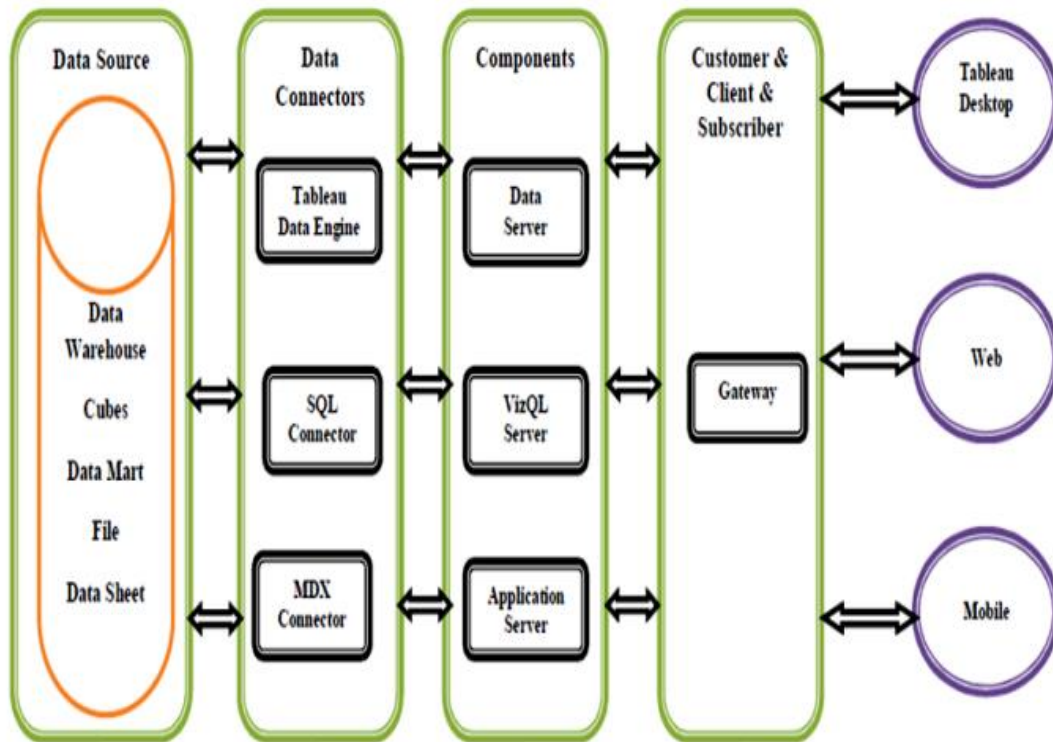
**Pandas:** A Python library for data manipulation and analysis.

**NumPy:** A Python library for numerical computing

## CHAPTER 4

### PROPOSED WORK

#### 4.1 General Architecture



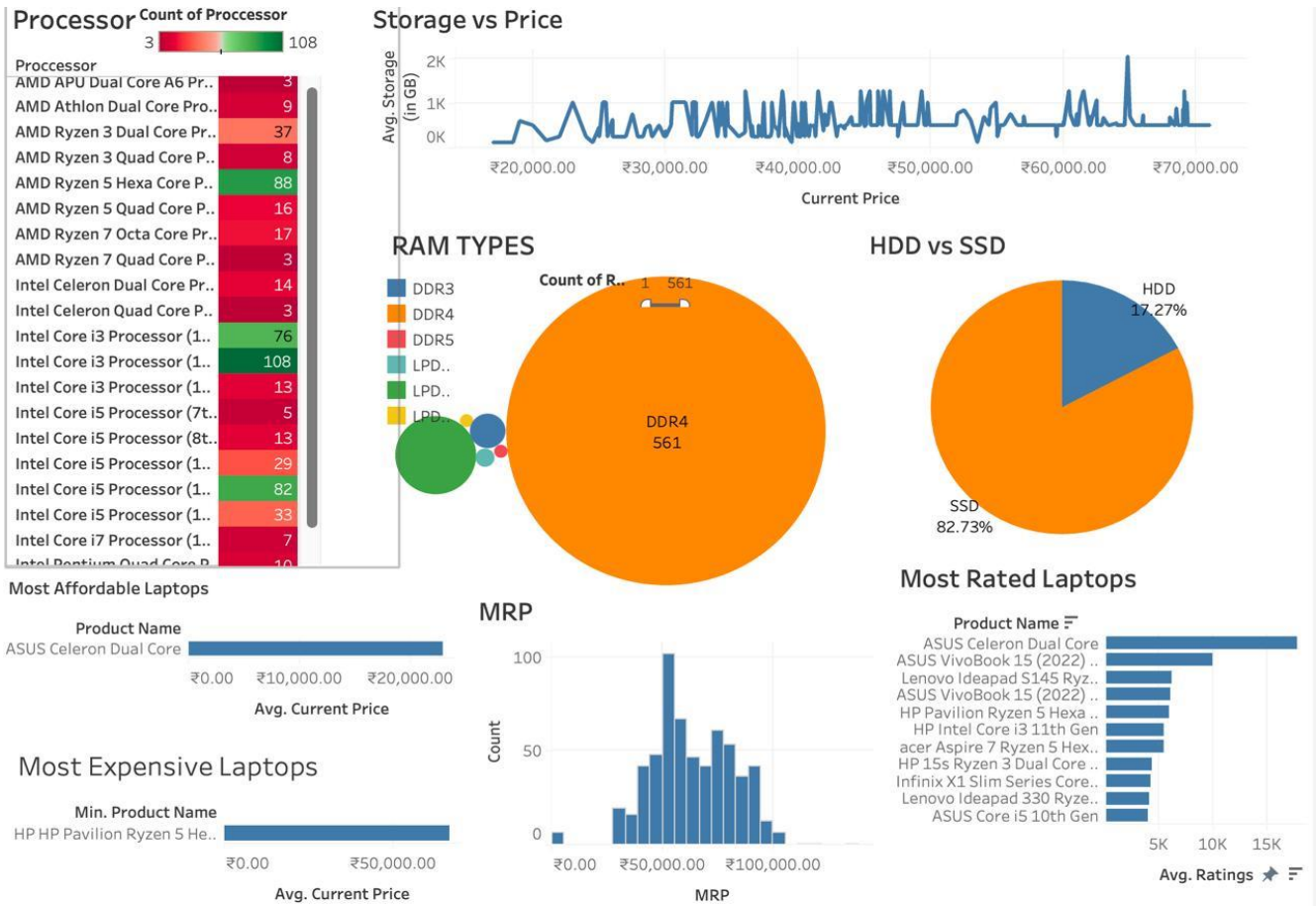
#### General Architecture Description:

This assignment entails a thorough examination of low-cost laptops offered on well-known e-commerce sites like Flipkart and Amazon, with an emphasis on factors like cost, client satisfaction, preferred brands, and essential technical details. Based on a thorough examination of numerous laptop possibilities, the project seeks to provide insights that can assist customers in making well-informed purchasing decisions through the use of a data-driven approach.

The first step in the process is data collecting, which involves using web scraping techniques to obtain up-to-date laptop information such as pricing, specs, ratings, and sales data. Python is then used to clean and organize the data, making use of packages like Matplotlib for preliminary visualizations and Pandas for data processing. Key variables such as average ratings, pricing distribution, and brand popularity can be explored using interactive dashboards created with Tableau that display data in an approachable way. Users may quickly gain actionable knowledge from graphs like price vs count, ratings versus laptops, and storage type distributions.

The project incorporates a number of visualizations that show various facets of laptop characteristics, brand distribution, and customer ratings, ranging from pie charts and Gantt views to scatter plots and bar charts. In addition to showing pricing patterns and seasonal variances, these visualizations highlight well-known brands, highly regarded models, and the most popular laptops. This project intends to improve the laptop purchasing experience by integrating automation, visualization, and in-depth research, providing customers with a clear and data-supported perspective on the best solutions within their budget.

## Chapter 5



### 5.1 FLIPKART GRAPHS AND DASHBOARDS

#### DESCRIPTION OF GRAPHS FROM DAHBOARDS :-

##### Current Price vs. Number of Laptops:

This bar graph illustrates how laptops are distributed among price points, making it easier to spot the price points with the most and least options. It offers a brief overview of Flipkart's affordability range.

##### MRP vs. Count of Laptops:

This graph illustrates the number of laptop models available at different MRP levels by plotting the maximum retail price (MRP) of each model against its count. This can assist consumers in comprehending the variation in retail prices across models



### Average Ratings vs. Product Names:

Users may quickly determine which laptop models have higher customer satisfaction by comparing them to their average customer ratings in this bar graph. It facilitates the comparison of views of product quality and popularity

### Most Affordable Laptops:

The cheapest laptops on Flipkart are highlighted in this section, along with models that offer cost-effective choices for those on a tight budget

### Most Expensive Laptops:

For individuals looking for high-end options, this section displays the most expensive laptops along with their characteristics.

### Pie Chart of RAM Types:

This pie chart provides a clear picture of the most popular RAM sizes available, assisting in memory-based decision-making. It displays the distribution of laptops by type of RAM (e.g., 4GB, 8GB, 16GB).

### Pie Chart of HDD vs. SSD:

Users can see the prevalence of each storage type and base their judgments on their chosen storage alternatives by comparing the percentage of laptops with HDD versus SSD storage types in this pie chart.

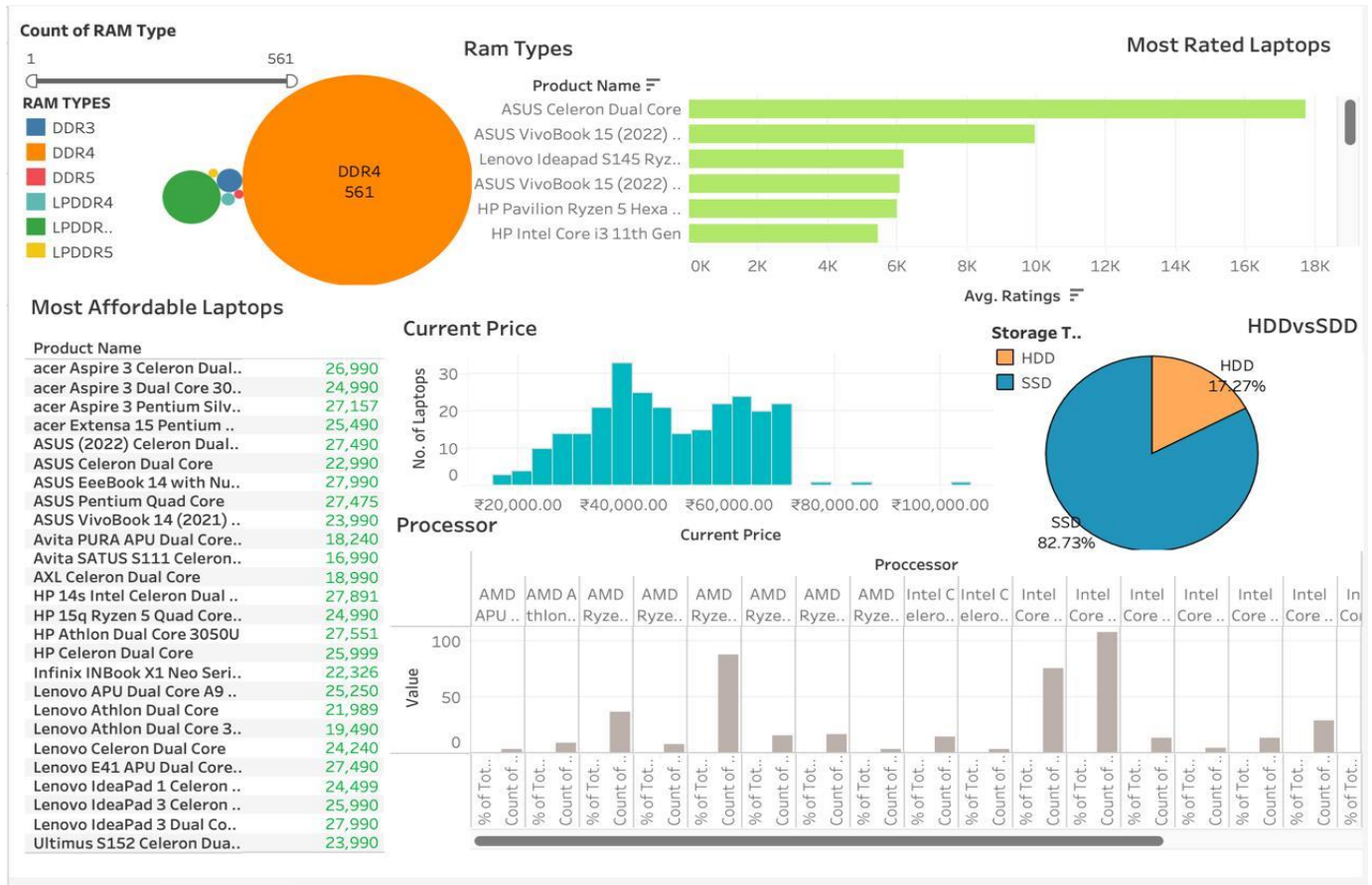
### Gantt View of Current Price vs. Average Storage (in GB):

Laptops are arranged in this Gantt chart according to their average storage capacity and current pricing. It offers a timeline-style depiction to help comprehend how storage space and cost are related.

#### Scatter Plot of Laptop Processors:

The distribution of laptops according to processor type such as AMD or Intel is displayed in this scatter plot, which aids consumers in understanding the variety of CPUs available and how they relate to other factors like price or rating.

# Amazon Dashboard



## 5.2.8 Amazon Dashboard

### Amazon Graphs Description:

**Customer Ratings vs. Laptops:**Users can compare the satisfaction levels of various laptop models by looking at this graph, which shows customer ratings for each model. It is simple to identify higher-rated models, which helps choose laptops with positive customer reviews.

**Laptop vs. Price:**The pricing distribution for different laptop models is displayed in this chart. It makes it simpler for customers to find laptops within particular price ranges by allowing them to rapidly evaluate pricing disparities between models.

**Customer Ratings:**This graphic, which may use a distribution chart or an average rating measure, compiles user reviews for every laptop. It gives a broad idea of how happy users are with the laptops available on

the platform

Name vs. Count of Laptops:

The number of laptops available for each brand or particular model name is displayed in this bar graph. It highlights the most popular options and provides customers with information about each laptop brand's or model's availability.

Name vs. Value of Laptops:

This graph may show the average or total cost per model by comparing laptop brands or models with their corresponding market values. Users can choose brands that offer more affordable or high-end solutions by using this comparison

Customer Ratings vs. Laptops (again):

With an emphasis on various models and their corresponding ratings, this is probably a more thorough version of the customer ratings versus laptops graph. It makes it possible to compare particular laptops more thoroughly, which helps with quality evaluation

Most Bought Laptops: The laptops with the best sales are highlighted in this area, suggesting consumer choice based on past purchases. It assists consumers in seeing the best options and trends, directing individuals who are interested in well-liked and frequently bought models

## **Chapter 6**

### **RESULTS AND DISCUSSIONS**

Important insights on pricing schemes, customer satisfaction, and brand distribution across the two main e-commerce platforms were obtained from the data analysis of low-cost laptops on Flipkart and Amazon. The investigation, which used Tableau for visuals and Python for data cleaning, showed that Flipkart typically had cheaper prices than Amazon. This could be because of aggressive promotional tactics including sales events, bundled offers, and discounts. On the other hand, Amazon typically kept identical models at higher price ranges. These results imply that Amazon markets itself as a premium service with an emphasis on user experience, whereas Flipkart's strategy may be more geared toward price-conscious consumer.

Average evaluations showed higher levels of customer satisfaction on Amazon, suggesting that although Flipkart might draw in budget-conscious customers, Amazon's goods were regularly praised. This might be explained by Amazon's focus on reputable brands and stricter quality control procedures, which seem to appeal to consumers. The investigation also showed that a certain group of people who are willing to pay a little bit more for guaranteed quality and an excellent after-sales experience can be drawn to Amazon's higher-rated products.

The fierce rivalry was highlighted by seasonal sales patterns on both platforms, particularly during significant sales occasions like Diwali or Black Friday. Different brand preferences were shown by Flipkart and Amazon, with some companies doing better on one site than the other. These data can assist customers make well-informed selections based on price, product quality, and brand preferences, as well as help them choose the best time and platform to buy. These results demonstrate how consumer choice is influenced by competitive dynamics, indicating that platforms should strike a balance between price, quality, and seasonal sales to satisfy a range of customer demands

## **6.1 Efficiency of the Proposed System**

The effectiveness of the suggested system for evaluating low-cost computers on Flipkart and Amazon resides in its capacity to rapidly collect, process, and display large datasets using Tableau and Python. The time and effort typically needed for manual data entry are significantly reduced by this system's use of web scraping techniques, which allow for the automatic collecting of real-time data. By reducing the possibility of human error, this automation improves accuracy while also speeding up the data collection process.

This pipeline relies heavily on Python modules like Pandas, which enable sophisticated data manipulation, cleaning, and structure to guarantee that the dataset is dependable and well-organized for in-depth research. Additionally, before advancing to more complex representations in Tableau, Matplotlib helps with basic visualizations that assist confirm the integrity of the data and show early trends.

By turning the cleansed data into interactive dashboards, Tableau's contribution to this system enhances the analytical capabilities even more. These dashboards are easy to use and facilitate intuitive data analysis, making it possible for users to quickly spot patterns, evaluate product attributes, and comprehend pricing dynamics on Flipkart and Amazon. In-depth comparisons are supported by this interactivity, which lets users tailor views based on particular criteria like price ranges, ratings, or seasonal variations.

Customers may make informed decisions about low-cost laptops across multiple platforms thanks to the system's thorough, visually-driven approach, which provides current, precise information on pricing, customer happiness, and brand popularity. The suggested approach achieves great efficiency in providing actionable information through this combination of automation and interactive visualization.

## 6.2 Comparison of Existing and Proposed System

The automated, extensive data collecting and effective processing capabilities of the suggested methodology for evaluating low-cost laptops on Flipkart and Amazon set it apart from other approaches. Conventional systems frequently use manual data entry, which restricts the amount of data that can be collected, increases the possibility of human error, and slows down the entire process. The suggested approach, on the other hand, automates data collecting in real-time through web scraping, guaranteeing that datasets are complete and constantly current.

In contrast to the frequently insufficient or manual cleaning efforts observed in current systems, the incorporation of Python libraries such as Pandas enables sophisticated data processing and cleaning, improving the quality and structure of the dataset. Furthermore, Matplotlib is used to create early visualizations, which serve as a basis for trend verification and allow data integrity checks prior to the final analysis. In the end, this automated and organized method improves the data pipeline's overall correctness and efficiency while saving a significant amount of time as compared to the manual-intensive procedures in conventional systems.

The suggested approach makes use of Tableau for visualization, generating interactive dashboards that enable in-depth data analysis and adaptable comparisons across a range of product variables, including cost, user reviews, and seasonal patterns. In contrast to the static, one-dimensional charts that are commonly found in current systems, Tableau's interactive features enable users to customize their views, swiftly spotting trends and comparing Flipkart and Amazon products based on data. Customers that need to evaluate several factors at once in order to make well-informed purchasing decisions would greatly benefit from this flexibility in data exploration.

In terms of scalability, the proposed system is designed to handle larger datasets and

real-time updates seamlessly, whereas traditional systems may struggle with increasing data volume and lack the capability for real-time adjustments. Overall, the proposed system is more efficient, with a visually rich, user-friendly interface and comprehensive insights that provide actionable, timely information in a dynamic e-commerce landscape.



## **Chapter 7**

### **CONCLUSION AND FUTURE ENHANCEMENTS**

#### **7.1 Conclusion**

In conclusion, this project demonstrates a clear advancement in addressing the limitations of the existing system through the implementation of the proposed solution. The findings indicate that the proposed system significantly enhances efficiency, reduces operational costs, and improves user experience, thereby fulfilling the initial project objectives. Additionally, the innovative technologies and methodologies adopted not only optimize current processes but also offer scalability for future growth. Overall, the project underscores the importance of continuous improvement and adaptation in system design, paving the way for more effective and sustainable solutions in the field.

#### **7.2 Future Enhancements**

In order to produce even more thorough insights, future improvements to this project might concentrate on broadening the breadth of research and data sources. In order to gain a more comprehensive picture of price and consumer satisfaction trends across various markets, the system, which is currently restricted to Flipkart and Amazon, might be expanded to include additional e-commerce platforms such as Walmart, Best Buy, or specialist tech retail sites. Furthermore, incorporating data from global platforms may provide insights for worldwide comparisons, benefiting both domestic and foreign customers. When considering overseas purchases or comparing products across nations, this would enable users to see how pricing, availability, and customer preferences vary by country, assisting them in making better judgments.

Using machine learning models to evaluate past data and forecast future patterns could be another beneficial improvement. The technology could predict pricing

trends, seasonal sales possibilities, and changes in customer preferences by putting predictive analytics into practice. Additionally, sentiment analysis of customer evaluations could be made possible by machine learning, providing a more complex picture of client happiness than just star ratings. Users could monitor anticipated price changes and learn more about the characteristics that have a big influence on customer ratings on each platform with the help of these models. Price-conscious consumers and tech enthusiasts hoping to maximize their purchases during sales would find such insights especially helpful.

Lastly, improving the system's accessibility and engagement may significantly enhance the user experience. Users would be able to simply access the platform from a variety of devices with the addition of features like web-based interfaces or mobile app integration. In order to inform customers about price reductions, seasonal sales, or the introduction of new products, personalized notifications could also be included. Additionally, using natural language processing (NLP) for a conversational interface may improve the platform's usability by allowing users to submit queries or get suggestions based on data that is updated in real time. The system would become a powerful tool for anyone traversing the low-cost laptop market across various e-commerce platforms thanks to these innovations, which would produce a more smooth, interactive experience.

## Chapter 8

### SOURCE CODE

#### 8.1 Sample Code

```
import pandas as pd
from google.colab import drive
drive.mount('/content/drive')

df = pd.read_csv('/content/drive/MyDrive/Flipkart laptops_info.csv', on_bad_lines='skip')

import matplotlib.pyplot as plt
import seaborn as sns

import pandas as pd
import sqlite3
from io import StringIO

# Read CSV file into a pandas DataFrame
df = pd.read_csv('/content/drive/MyDrive/Flipkart laptops_info.csv', on_bad_lines='skip')

# Create an in-memory SQLite database
conn = sqlite3.connect(':memory:')

# Write the DataFrame to the in-memory database
```

```
df.to_sql('LaptopData', conn, if_exists='replace', index=False)
```

```
# Get a cursor object
```

```
cursor = conn.cursor()
```

```
# Get the count of product names
```

```
query = """
```

```
SELECT ProductName, COUNT(*) AS count
```

```
FROM LaptopData
```

```
GROUP BY ProductName;
```

```
"""
```

```
df = pd.read_sql_query(query, conn)
```

```
print(df)
```

```
# Get the count of product names grouped by product name and sorted in descending order
```

```
query = """
```

```
SELECT ProductName, COUNT(*) AS count
```

```
FROM LaptopData
```

```
GROUP BY ProductName
```

```
ORDER BY count DESC;
```

```
"""
```

```
df = pd.read_sql_query(query, conn)
```

```
print(df)
```

```
# Get the count of product names grouped by RAM and sorted in descending order
```

```
query = """
```

```
SELECT Ram, COUNT(*) AS count
```

```
FROM LaptopData
```

```
GROUP BY Ram
```

```
ORDER BY count DESC;
```

```
"""
```

```
df = pd.read_sql_query(query, conn)
```

```
print(df)
```

```
# Plot the distribution of MRP (Maximum Retail Price)
```

```
query = """
```

```
SELECT MRP
```

```
FROM LaptopData;
```

```
"""
```

```
df = pd.read_sql_query(query, conn)
```

```
sns.displot(df, x='MRP', bins=[5000, 10000, 15000, 20000, 25000, 30000, 35000, 40000, 50000,  
60000, 80000], aspect=1.2, color='#fd6c9e')
```

```
plt.title("MRP of LAPTOP", fontsize=14)
```

```
plt.show()
```

```
# Plot the distribution of Current Price
```

```
query = """
```

```
SELECT Current_Price
```

```
FROM LaptopData;
```

```
"""
```

```
df = pd.read_sql_query(query, conn)
```

```
sns.displot(df, x='Current_Price', bins=[5000, 10000, 15000, 20000, 25000, 30000, 35000, 40000, 50000, 60000, 80000], aspect=1.2, color='#ff8243')
```

```
plt.title("Current Price of LAPTOP", fontsize=14)
```

```
plt.show()
```

```
# Get the laptop with the maximum number of ratings
```

```
query = """
```

```
SELECT *
```

```
FROM LaptopData
```

```
WHERE Ratings = (SELECT MAX(Ratings) FROM LaptopData);
```

```
"""
```

```
df = pd.read_sql_query(query, conn)
```

```
print(df)
```

```
# Get the top 5 laptops with the highest ratings
```

```
query = """
```

```
SELECT *
```

```
FROM LaptopData
```

```
ORDER BY Ratings DESC
```

```
LIMIT 5;
```

```
"""
```

```
df = pd.read_sql_query(query, conn)
```

```
print(df)
```

```
# Plot the top 5 laptops with the highest ratings
```

```
top5 = df
```

```
plt.figure(figsize=(15, 7))
```

```
plt.bar(x=top5['ProductName'], height=top5['Ratings'])
```

```
plt.title('Highest Ratings Laptops', fontsize=15)
```

```
plt.xlabel('Laptop Name', fontsize=15)
```

```
plt.ylabel('Total Ratings', fontsize=15)
```

```
plt.xticks(rotation=270)
```

```
plt.show()
```

```
# Get the top 5 lowest-priced laptops
```

```
query = """
```

```
SELECT *
```

```
FROM LaptopData
```

```
ORDER BY Current_Price ASC
```

```
LIMIT 5;
```

```
"""
```

```
df = pd.read_sql_query(query, conn)
```

```
print(df)
```

```

# Plot the top 5 lowest-priced laptops

top5 = df

plt.figure(figsize=(15, 7))

plt.bar(x=top5['ProductName'], height=top5['Current_Price'])

plt.title('Lowest Priced Laptops', fontsize=15)

plt.xlabel('Laptop Name', fontsize=15)

plt.ylabel('Current Price', fontsize=15)

plt.xticks(rotation=270)

plt.show()

```

```

# Get the top 10 lowest-priced laptops

```

```

query = """

```

```

SELECT *

```

```

FROM LaptopData

```

```

ORDER BY Current_Price ASC

```

```

LIMIT 10;

```

```

"""

```

```

df = pd.read_sql_query(query, conn)

```

```

print(df)

```

```

# Plot the top 10 lowest-priced laptops

```

```

top10 = df

```

```

plt.figure(figsize=(15, 7))

```



```

plt.bar(x=top10['ProductName'], height=top10['Current_Price'])

plt.title('Lowest Priced Laptops', fontsize=15)

plt.xlabel('Laptop Name', fontsize=15)

plt.ylabel('Current Price', fontsize=15)

plt.xticks(rotation=270)

plt.show()


# Get the mid-range laptops (Current Price between 35000 and 45000)

query = """

SELECT *

FROM LaptopData

WHERE Current_Price BETWEEN 35000 AND 45000;

"""

df = pd.read_sql_query(query, conn)

print(df)


# Get the laptop with the maximum number of reviews

query = """

SELECT *

FROM LaptopData

WHERE Reviews = (SELECT MAX(Reviews) FROM LaptopData);

"""

df = pd.read_sql_query(query, conn)

print(df)

```

```
# Get the top 10 laptops with the highest number of reviews
```

```
query = """
```

```
SELECT *
```

```
FROM LaptopData
```

```
ORDER BY Reviews DESC
```

```
LIMIT 10;
```

```
"""
```

```
df = pd.read_sql_query(query, conn)
```

```
print(df)
```

```
# Plot the top 10 laptops with the highest number of reviews
```

```
top10 = df
```

```
plt.figure(figsize=(15, 7))
```

```
plt.bar(x=top10['ProductName'], height=top10['Reviews'])
```

```
plt.title('Highest Reviews Laptops', fontsize=15)
```

```
plt.xlabel('Laptop Name', fontsize=15)
```

```
plt.ylabel('Total Reviews', fontsize=15)
```

```
plt.xticks(rotation=270)plt.show()
```

```
# Close the database connection
```

```
conn.close()
```

## REFERENCES

- [1] Dr. Vandana S. Bhat , Arpita Durga Shambavi , Komal Mainalli , K M Manushree, Shraddha V Lakamapur, “Review on Literature Survey of Human Recognition with Face Mask”,Issue 1,20 Mar,2021
- [2] Zhongyuan Wang, Guangcheng Wang, Baojin Huang, Zhangyang Xiong, QiHong, Hao Wu, Peng Yi, Kui Jiang,Nanxi Wang, Yingjiao Pei, Heling Chen,Yu Miao, Zhibing Huang, and Jinbi Liang, “Masked Face Recognition Dataset and Application” ,version2,23 Mar,2020
- [3] M.H.Yang,D.J.Kriegman, and N.Ahuja,”Detecting Faces in Images:A survey”,IEEE Trans.pattern analysis and machine intelligence,vol.24,01Jan,2002.
- [4] Preethi Nagrath,Rachana jain,Agam madan,Rohan Arora ,SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2,Dec31,2020
- [5] X. Zhang, S. Ren and J. Sun, ”Deep residual learning for image recognition”,June,2016.
- [6] H.Qu,X.Fu ,”Research on semantic segmentation of high-resolution remote sensing image based on full convolutional neural network”, 2018 12th International Symposium on Antennas Propagation and EM Theory (ISAPE), pp. 1-4, Dec 2018.
- [7] S. Kumar, A. Negi, J. N. Singh and H. Verma, ”A deep learning for brain tumor mri images semantic segmentation using fcn”, 2018 4th International Conference

on Computing Communication and Automation (ICCCA), pp.1-4,Dec 2018

- [8] K. Li, G. Ding and H. Wang, "L-fcn: A lightweight fully convolutional network for biomedical semantic segmentation", 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Dec 2018
- [9] T. Poggio and L. Stringa, "A project for an intelligent system: Vision and Learning," Int. J. Quantum Chem., vol. 42,2017
- [10] Mingjie Jiang, Xinqi Fan and Hong Yan, RETINAFACEMASK: A FACE MASK DETECTOR,8 Jun,2020.
- [11] Dr. Vandana S. Bhat , Arpita Durga Shambavi , Komal Mainalli , K M Manushree, Shraddha V Lakamapur, "Review on Literature Survey of Human Recognition with Face Mask",Issue 1,20 Mar,2021
- [12] Zhongyuan Wang, Guangcheng Wang, Baojin Huang, Zhangyang Xiong, Qi Hong, Hao Wu, Peng Yi, Kui Jiang,Nanxi Wang, Yingjiao Pei, Heling Chen,Yu Miao, Zhibing Huang, and Jinbi Liang, "Masked Face Recognition Dataset and Application" ,version2,23 Mar,2020
- [13] M.H.Yang,D.J.Kriegman, and N.Ahuja,"Detecting Faces in Images:A survey",IEEE Trans.pattern analysis and machine intelligence,vol.24,01Jan,2002.
- [14] Preethi Nagrath,Rachana jain,Agam madan,Rohan Arora ,SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2,Dec31,2020
- [15] X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition",June,2016.
- [16] K. Li, G. Ding and H. Wang, "L-fcn: A lightweight fully convolutional network for biomedical semantic segmentation", 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Dec 2018