

Laptop Price Analysis

Data Exploration & Visualization

- Workflow:
 1. Data Exploration
 2. Data Cleaning
 3. Data Visualization
 4. Hypothesis Testing

Presented By: Naresh Kumar. C

Milestone: Python Data Exploration Project

Date: December 24, 2024



Introduction



Objective:

Understand and analyze the factors influencing laptop prices.



Dataset:

A CSV file containing details of various laptops, including price, RAM, weight, OS, and more.



Importance:

- Data understanding and preprocessing are critical for meaningful analysis.
- Visualizations help uncover patterns and relationships in the data.

Data Understanding

Dataset Summary

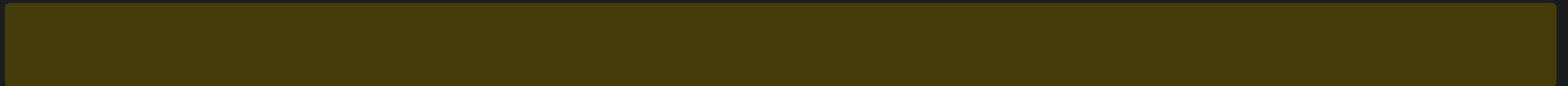
Rows : 1275

Columns : 23

Data Types:- Float : 4 Columns

Int : 5 Columns

Object : 14 Columns



Initial Exploration

Patterns Identified:

Missing Values

No Missing values.

Outliers

No outlier was found in numerical features.

Categorical Analysis

Data types and unique values



Missing Value Handling

Techniques used :

- Missing values identified using `isnull()`.
- No Missing values.

Outlier Handling

Detection Method:

- Box plot method to visualize outliers in numerical feature.

Impact:

- No outlier was found.
-

Fixing Rows and Columns

Examples Fixes

- Replaced 24GB RAM with 16GB.
- Replaced 64GB RAM with 32GB.
- Replaced "No OS" with "Windows 10".

Comparison

Shape Before & Shape After:

Fixing Rows and Columns

[47]: data[data["Ram"] == 24]

	Company	Product	TypeName	Inches	Ram	OS	Weight	Price_euros	Screen	ScreenW	...	RetinaDisplay	CPU_company	CPU_freq	CPU_model	Primary
438	Lenovo	ThinkPad T470s	Ultrabook	14.0	24	Windows 10	1.32	2382.0	Standard	2560	...	No	Intel	2.7	Core i7 7500U	
494	Asus	Rog G752VL-LH711T	Gaming	17.3	24	Windows 10	4.33	1269.0	Full HD	1920	...	No	Intel	2.6	Core i7 6700HQ	
517	Asus	ROG Zephyrus	Gaming	15.6	24	Windows 10	2.24	2968.0	Full HD	1920	...	No	Intel	2.8	Core i7 7700HQ	

3 rows x 23 columns

[48]: data["Ram"] = data["Ram"].replace(24,16)

[49]: data[data["Ram"] == 16]

	Company	Product	TypeName	Inches	Ram	OS	Weight	Price_euros	Screen	ScreenW	...	RetinaDisplay	CPU_company	CPU_freq	CPU_model	Primary
1066	Asus	ROG G701VO	Gaming	17.3	64	Windows 10	3.58	3975.0	Full HD	1920	...	No	Intel	2.7	Core i7 6820HK	

1 rows x 23 columns

[50]: data["Ram"] = data["Ram"].replace(64,32)

Made with Gamma

Handling Invalid Values

Identified Issues:

- Erroneous entries in RAM and OS.

Actions Taken:

- Replaced invalid values with corrected entries.

Data Visualization (EDA)

1 Univariate Analysis

Histogram :

- Shows the distribution of prices with a density plot.

Countplot :

- Visualizes the count of laptops for each company.

Piechart :

- Displays the proportion of different RAM sizes.

BarPlot :

- Compares the CPU brand by Touchscreen

2 Bivariate Analysis

Scatter plots :

- Explores the relationship between laptop weight and price.

Bar plot :

- Compares the average price by company, category and different operating systems.

Factor plot :

- Combines TypeName and Touchscreen with average prices to analyze categorical relationships.

1 Multivariate Analysis

Pairplot:

- This plots pairwise relationships between the Price, Weight, and Screen Width.

Heatmap:

- Correlation matrix to study numerical feature interactions.

Hypothesis Testing

Chi-Square Test: To check if Touchscreen and IPS Panel are independent

Touchscreen vs IPS Panel

- Result:

- Chi2 Statistic Value :

24.02118831026901

- P-value:

13.692931228609764

- Conclusion :

Variables are not independent

T-Test: To compare prices between laptops with and without touchscreens:

Price vs Touchscreen

- Result:
 - T-Statistic:

Nan

- P-value:

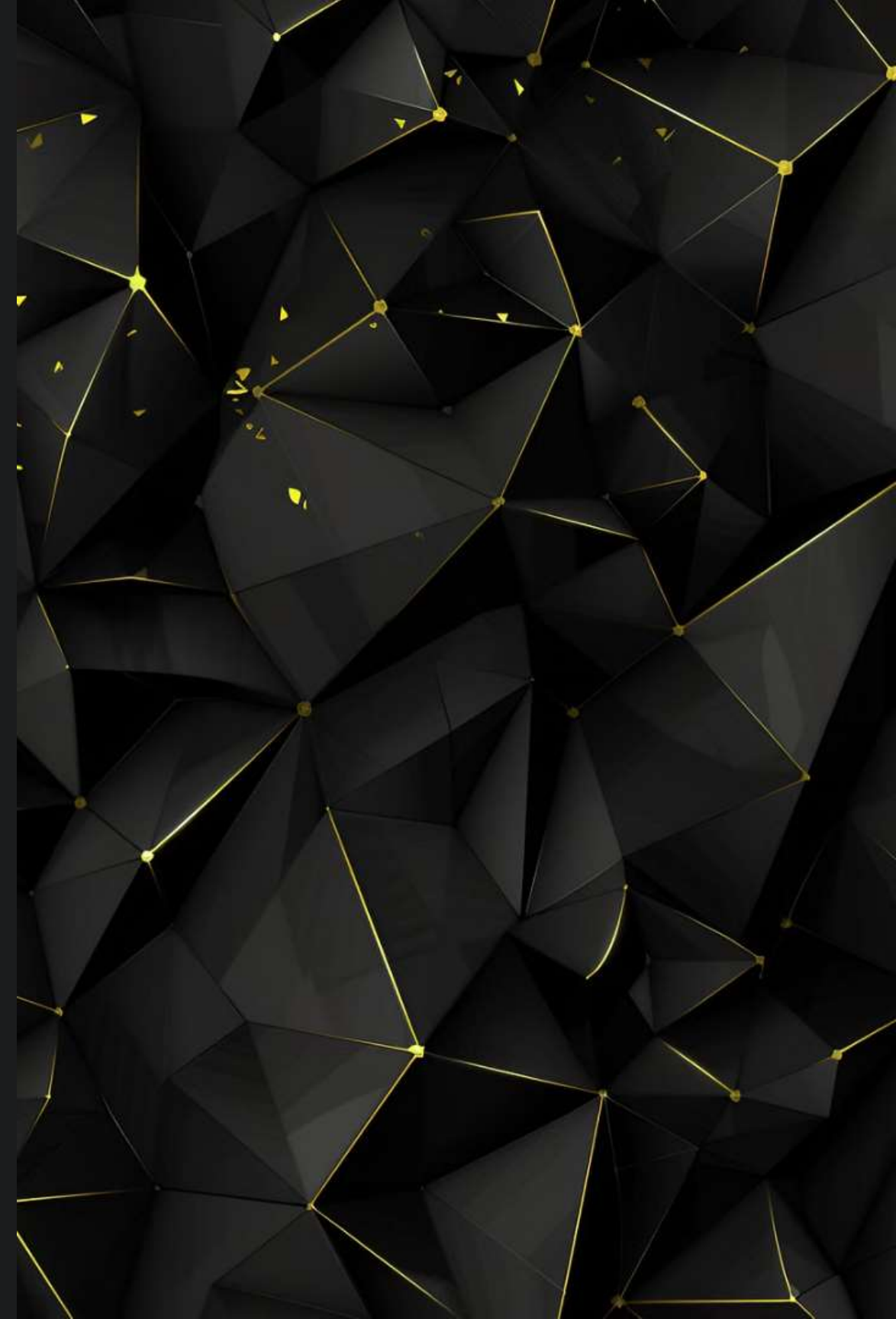
Nan

- Conclusion:

There is no significant difference between touchscreen and price.

Conclusion

- Summary of Findings :
 - Price influenced by RAM, weight, and brand.
 - Touchscreen availability impacts price significantly.



Thank You

