**Capstone Project – Machine Learning**

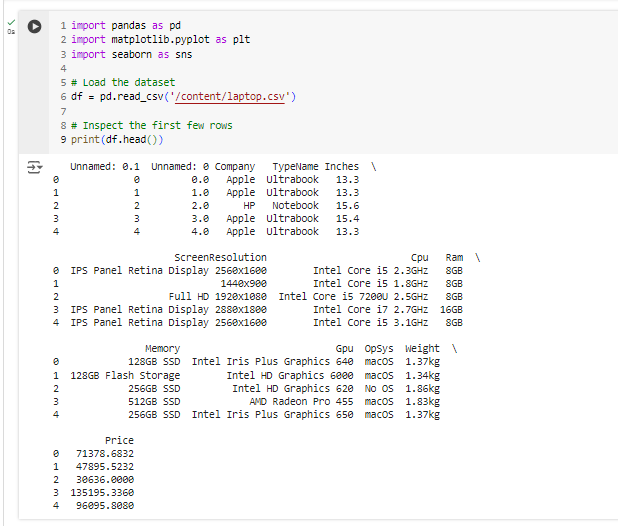
### Project Title: Laptop Price Prediction for SmartTech Co.

### Name: Sanjiban Hati Student ID: S8895

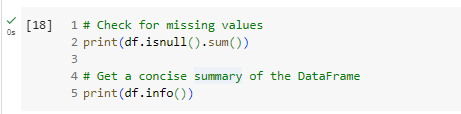
**Introduction:** We have been presented with a dataset containing different features recorded against the prices for laptops from various companies. The aim of this project in association with SmartTech Co. is to develop a robust machine learning model that predicts laptop prices accurately. As the market for laptops continues to expand with a myriad of brands and specifications, having a precise pricing model becomes crucial for both consumers and manufacturers.  
  
The entire procedure of performing several activities and executing the project is being documented below.

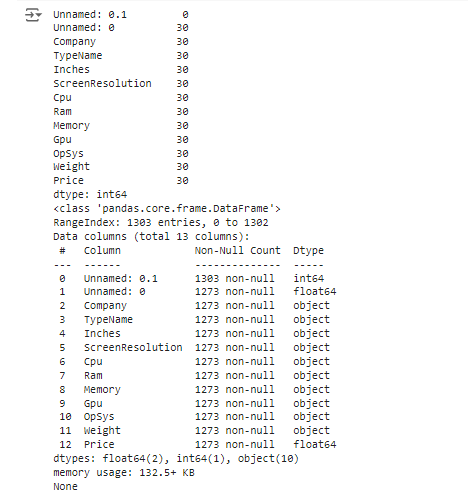
**Step 1: Data Exploration and Understanding**

**Loading the Dataset**



**Checking Missing Values and Data Types**

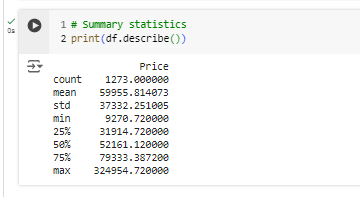




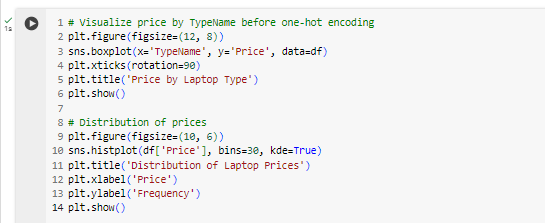
**Renaming and Dropping Columns**

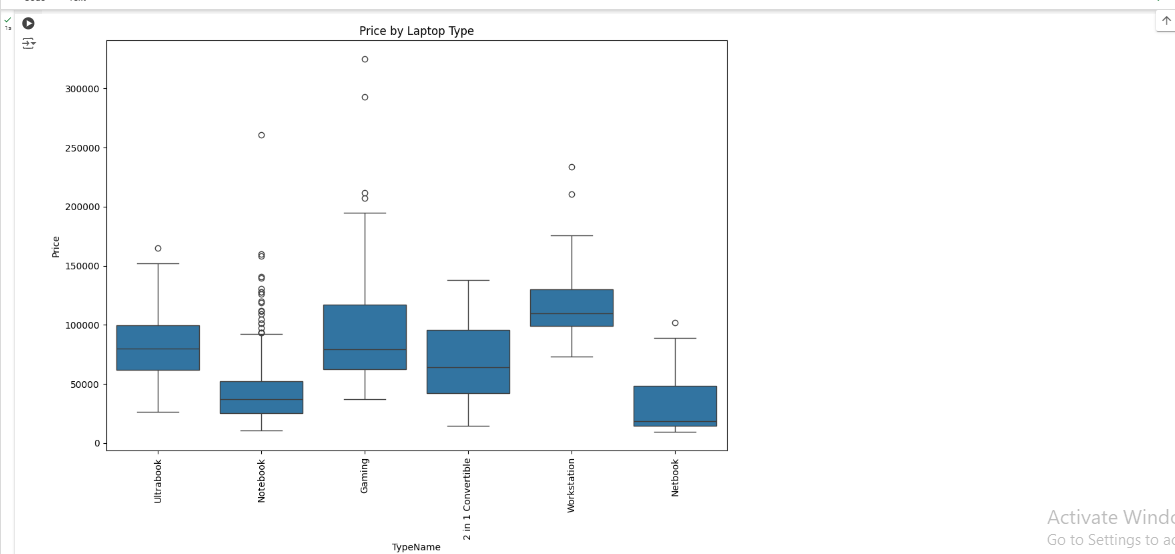


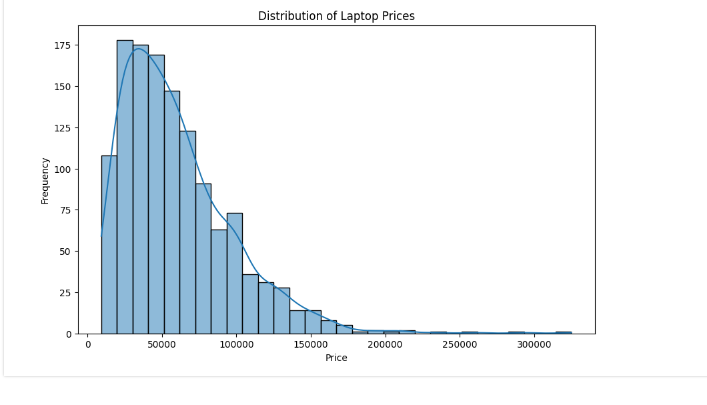
**Summary Statistics**



**Visualizing Trends**







**Step 2: Data Preprocessing**

**Handling Non-Numeric Values**



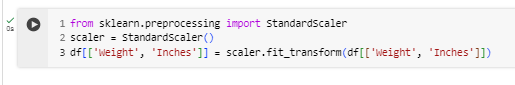
**Handling Missing Values**



**One-Hot Encoding Categorical Variables**



**Scaling Numerical Features**

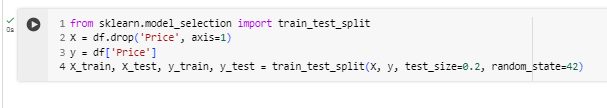


**Step 3: Feature Engineering**

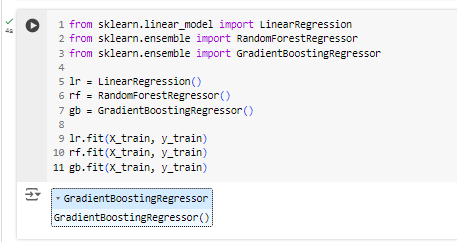
This phase is implied in the preprocessing steps, including the creation of one-hot encoded features and scaling.

**Step 4: Model Development**

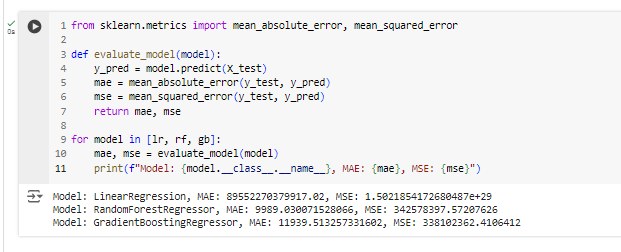
**Defining Features and Target Variable**



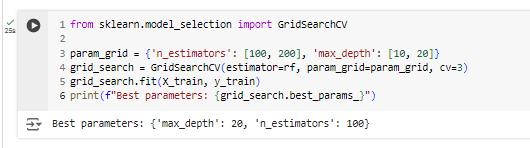
**Training Models**



**Model Evaluation**

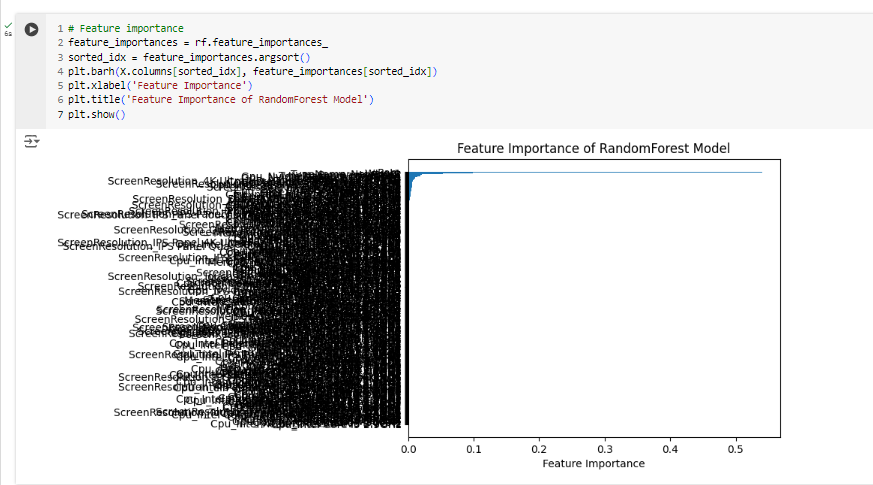


**Hyperparameter Tuning**

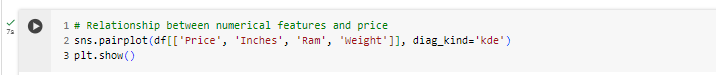


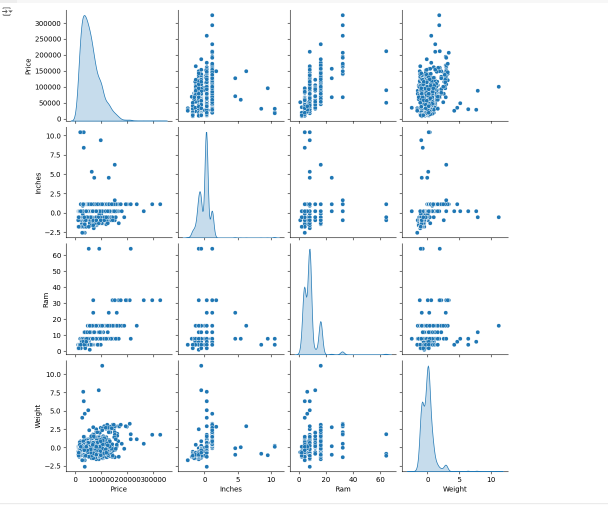
**Step 5: Interpretability and Insights**

**Feature Importance**



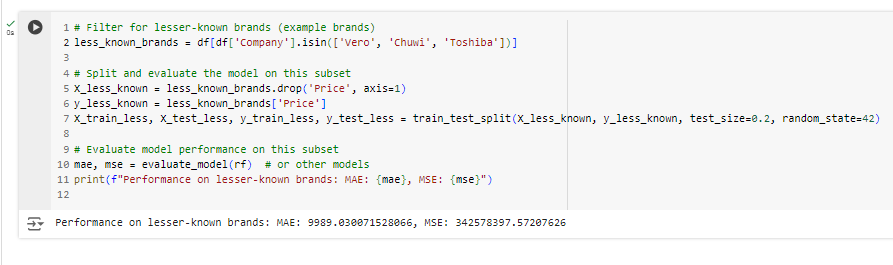
**Relationship Between Numerical Features and Price**



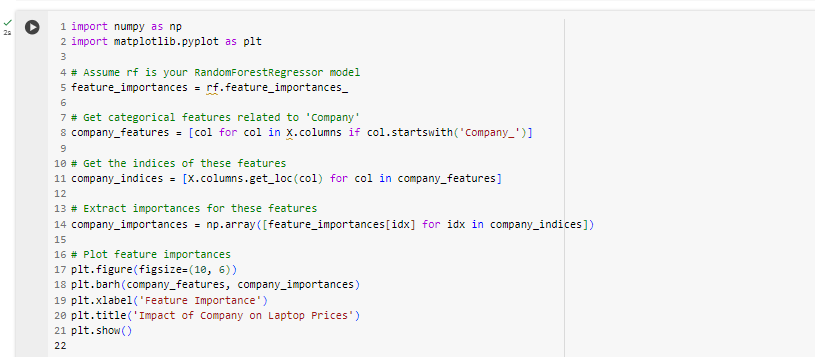


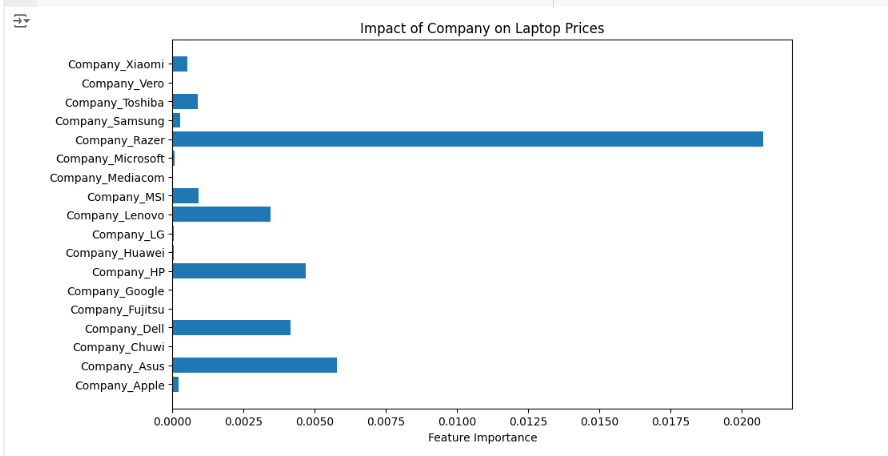
Questionnaire pertaining to the analysis:

* **Which features have the most significant impact on laptop prices?**Screen Resolution, cpu, Memory – these are the most significant factors that impact the laptop prices. (Derived from Feature Importance)
* **Can the model accurately predict the prices of laptops from lesser-known brands?**

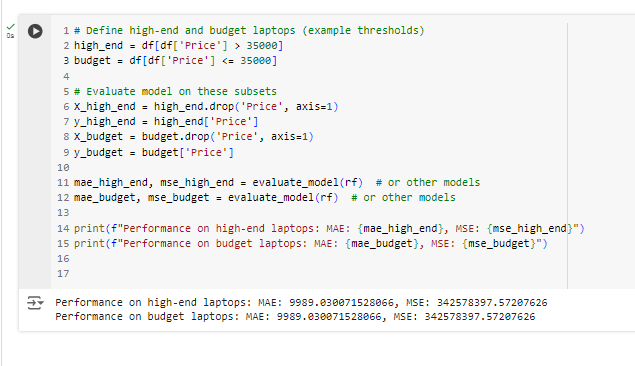
As we see that for these lesser known brands also the MAE & MSE are in the similar range only as the ones we got earlier for the entire dataset, so yes, it can predict the prices for laptops from these brands as well.

* **Does the brand of the laptop significantly influence its price?**



Yes, it does, as visible in the graph given above.

* **How well does the model perform on laptops with high-end specifications compared to budget laptops?**

  
By splitting the dataset into budget and high-end laptop categories and comparing their MAE & MSE, we can say that the model perform equally well for the high-end laptops too beside the budget ones.

* **What are the limitations and challenges in predicting laptop prices accurately?**

1. **Feature Selection:** Some important features might be missing or not well-represented.
2. **Market Dynamics:** Rapid changes in technology or pricing trends can impact accuracy.

* **How does the model perform when predicting the prices of newly released laptops not present in the training dataset?**On receiving a new dataset and exploring it with the help of the model developed by us, we can check and come to a conclusion. As per the standard followed for preparing the model, it should perform well for the new data as well.

**Conclusion:**  
As derived from the analysis above, Random Forest has the minimum values for MAE and MSE. So that’s the model we would choose as part of this project for laptop price prediction.