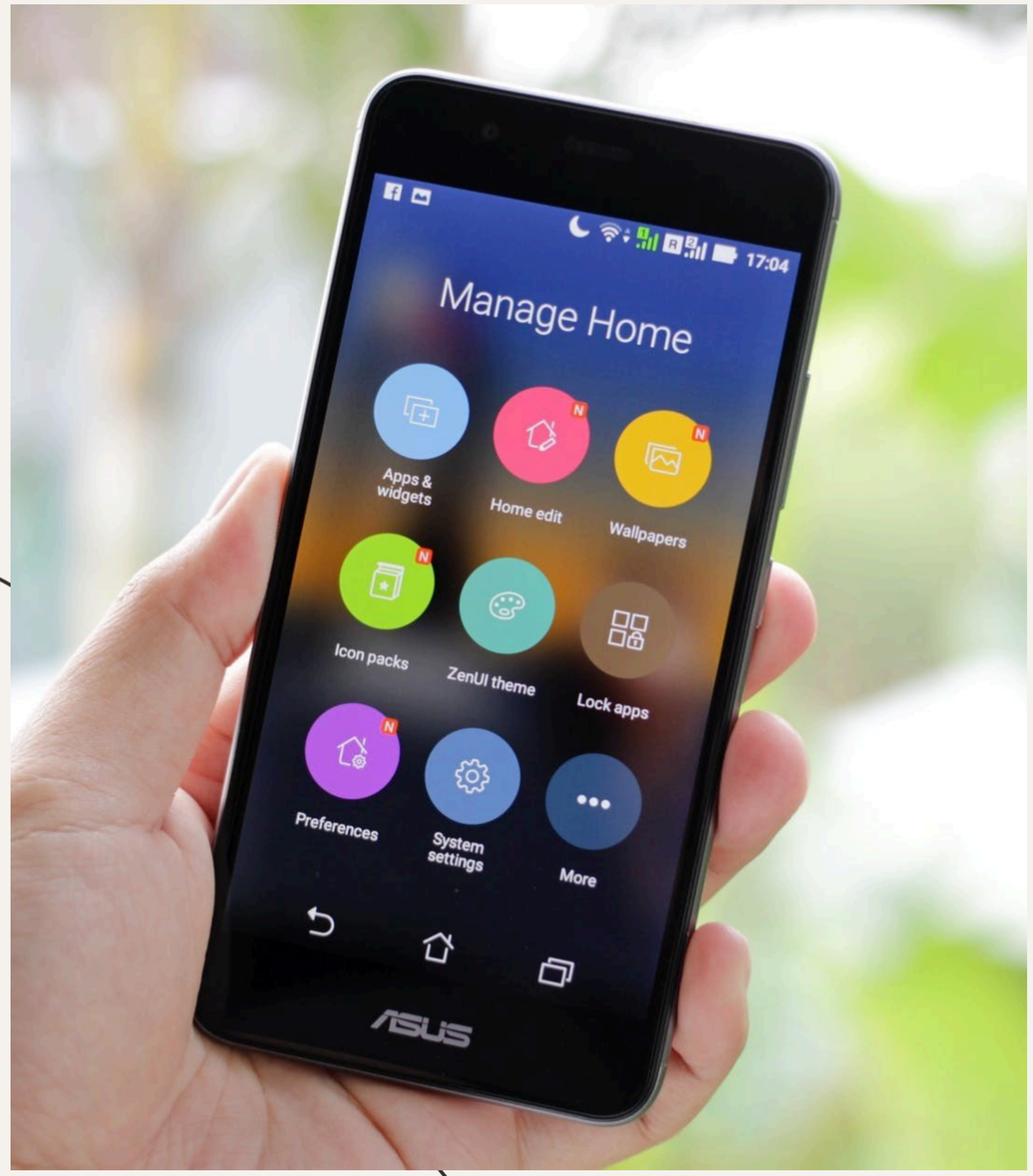
Project Title :- Mobile Price Prediction Model

Name :- AKSHAT JAIN

Date :- 1 August, 2024



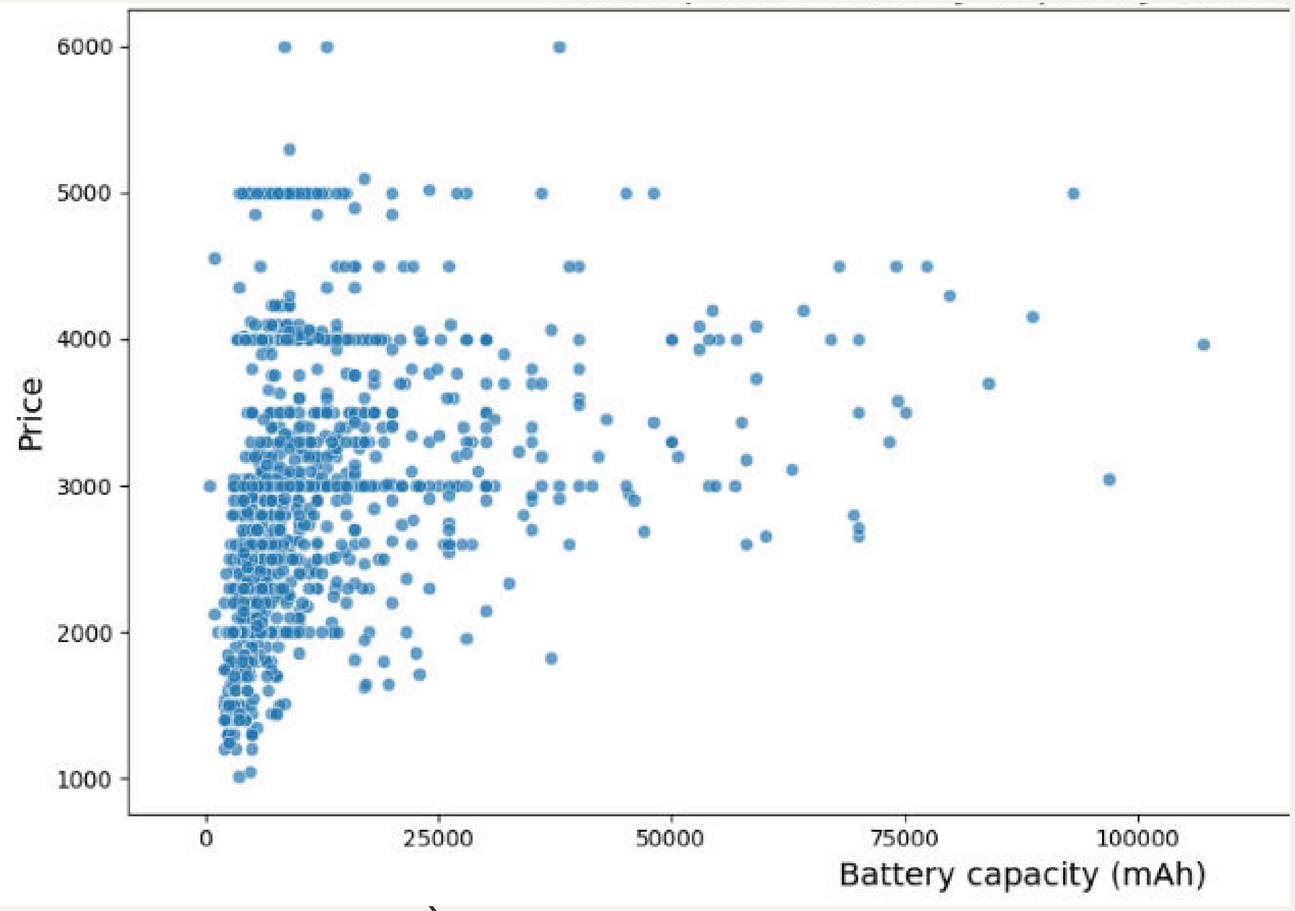
Objective :-

Develop a linear regression model to predict the price of mobile phones based on various features such as brand, name, cameras and market trends. The aim is to create an accurate pricing model that can estimate mobile phone prices effectively

Exploratory Data Analysis

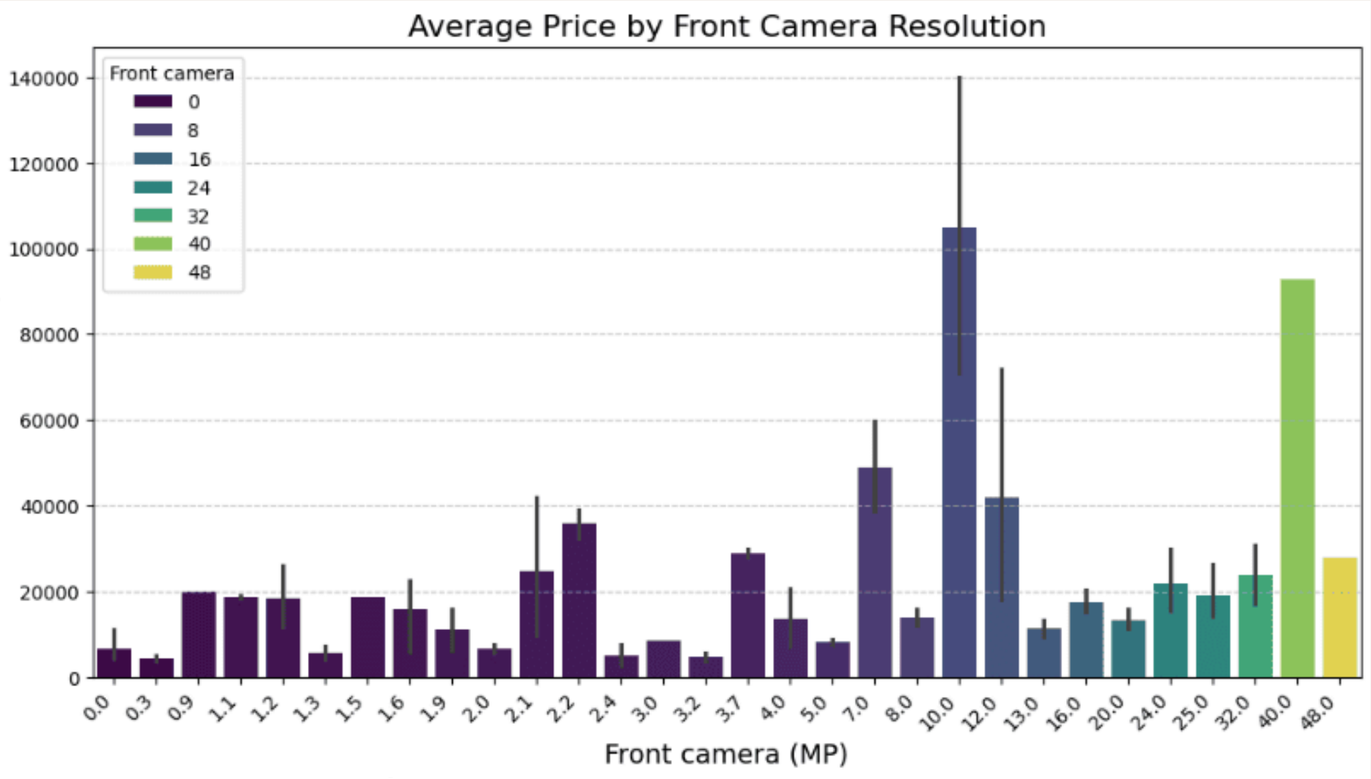
Identifying the right **data sources** is crucial for a successful machine learning project. We will discuss various types of **data**, including structured and unstructured, and how to evaluate their **quality** and **relevance** for your specific objectives.

BATTERY CAPACITY AND PRICING



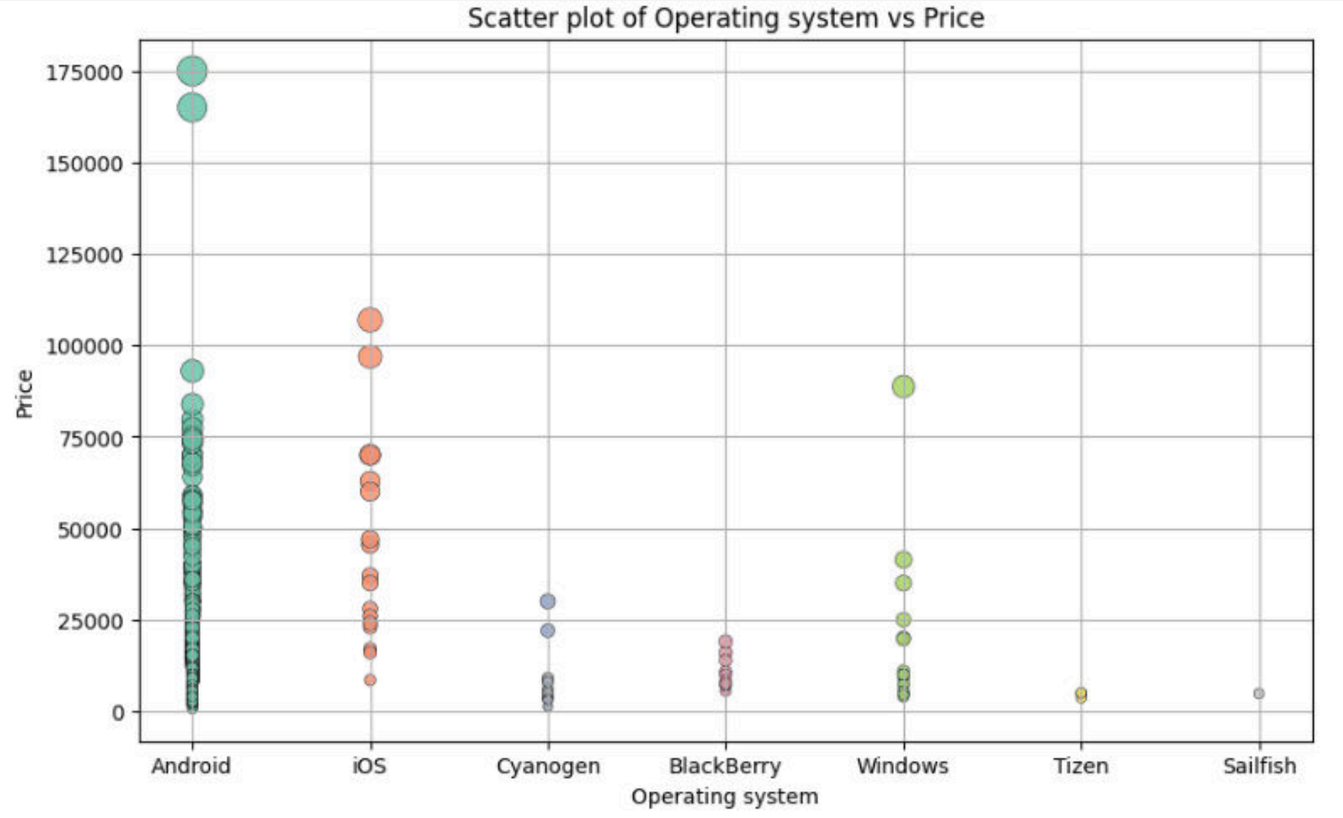
Phones with larger **battery** capacities tend to be priced **higher**. This is because extended battery life, which reduces the need for frequent **recharging**, adds value and justifies a higher price point.

CAMERA AND PRICING



The price of a Phone isn't always directly related to camera. For instance, mid-range phones might offer better camera for a better viewing experience, while high-end phone often prioritize a more compact design, which doesn’t always mean a better camera.

OPERATING SYSTEM TRENDS



Android leads as the predominant operating system, reflecting a vast and varied market with numerous brands and versions within the Android ecosystem.

Data Preprocessing Techniques

Before modeling, data must undergo **preprocessing** ensure accuracy and consistency. We will cover essential techniques like **Lable Encoding** and **OneHotEncoding for Catagorical values** to prepare your data for effective analysis and model training.

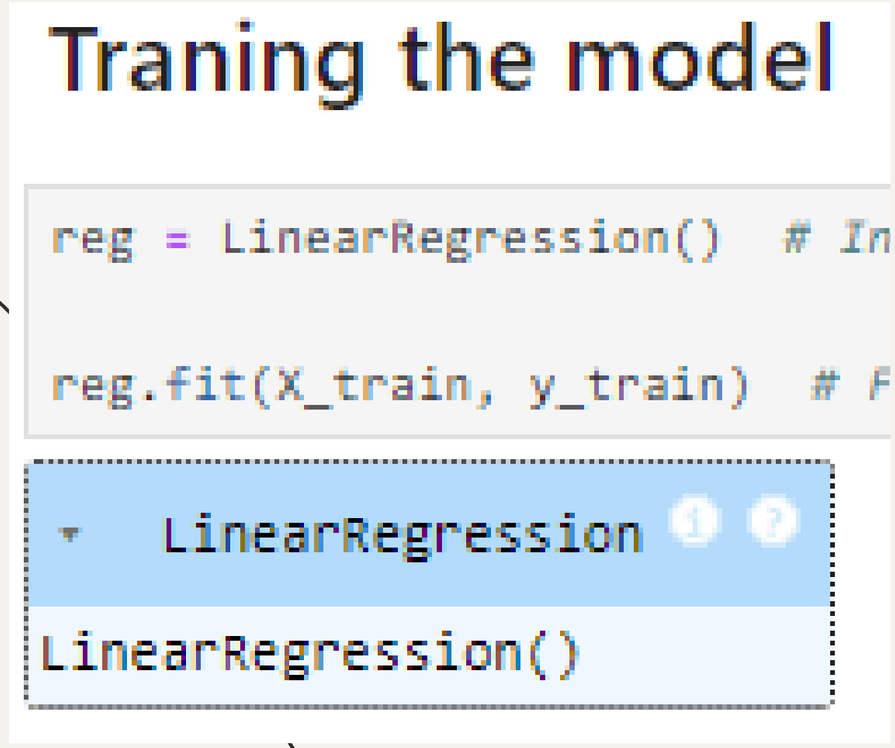
```
le = LabelEncoder()

for column in ['Brand', 'Model', 'Operating system']:
    le = LabelEncoder()
    x[column] = le.fit_transform(x[column])
```

```
ct = ColumnTransformer(
    transformers=[('encoder', OneHotEncoder(), [4, 13, 14, 15, 17, 18])],
    remainder='passthrough'
)

x = np.array(ct.fit_transform(x))
```

Model Building



Develop a predictive model for Mobile Price using **Linear Regression**. Although **KNN** was initially tested, it did not yield satisfactory accuracy. Linear Regression was chosen for its efficiency in handling binary outcomes and its ability to provide clear probability estimates, leading to improved model performance in this context.

Model Evaluation

```
print(f"Root Mean Squared Error (RMSE): {RMSE:.4f}")
print(f"Mean Absolute Error (MAE): {MAE:.4f}")
print(f"R-squared (R²): {R_squared:.4f}")
```

```
Root Mean Squared Error (RMSE): 7653.8744
Mean Absolute Error (MAE): 4855.3223
R-squared (R²): 0.7181
```

Conclusion

The model evaluation reveals a Root Mean Squared Error (RMSE) of 7653.8744 and a Mean Absolute Error (MAE) of 4855.3223, suggesting that while there is some prediction error, the model is reasonably accurate. The R-squared (R^2) value of 0.7181 indicates that approximately **72%** of the variance in the data is explained by the model, demonstrating a good fit and effective predictive performance. Overall, the model provides a solid basis for predictions, though there is room for improvement.

Thanks!