The background features a large white circle in the center, which is partially overlaid by a dark blue shape at the bottom. To the left of the circle is a light blue vertical rectangle, and to the right is a light pink vertical rectangle.

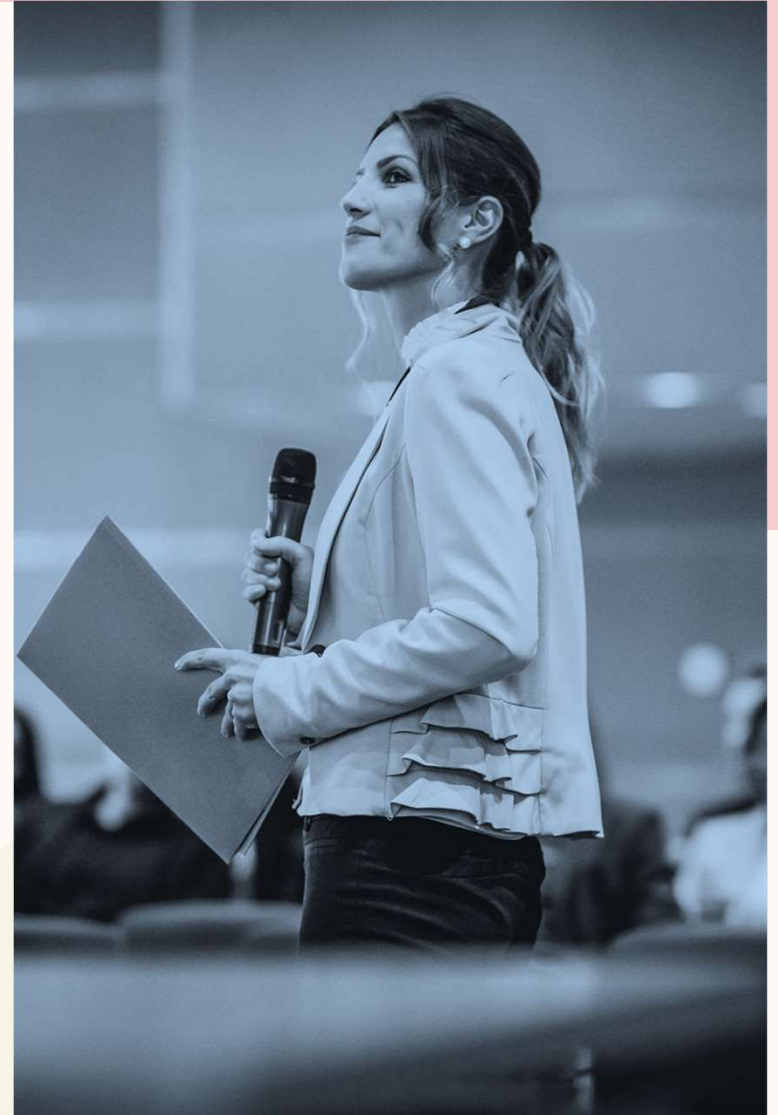
FEATURE EXTRACTION AND PRICE PREDICTION FOR MOBILE PHONES

INTRODUCTION

In this project, we have worked with a dataset that contains detailed information about various mobile phones, including their model, color, memory, RAM, battery capacity, rear camera specifications, front camera specifications, presence of AI lens, mobile height, processor, and, most importantly, the price. Our primary goal is to develop a predictive model for mobile phone prices.

POINTS DISCUSS

- Data description and summary
- Exploratory data analysis
- Heat map
- Machine learning algorithms 1. Logistic regression 2. Decision tree 3. Random forest classifier 4. Xgboost classifier
- conclusion





DATA PREPROCESSING

- **Data Cleaning:** Conversion of the "Price" column from string to numerical format, ensuring all commas are removed for proper numeric representation.
- **Handling Missing Values:** If any missing values are present, appropriate strategies like imputation or removal are applied.
- **Feature Encoding:** One-hot encoding is applied to categorical features such as Model, Color, Rear Camera, Front Camera, and Processor to convert them into a format suitable for machine learning algorithms.

EXPLORATORY DATA ANALYSIS

Distribution of Feature: Analyzing the distribution of numerical features like Memory, RAM, Battery, and Price to understand their range and central tendencies.

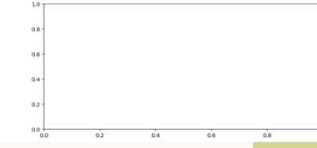
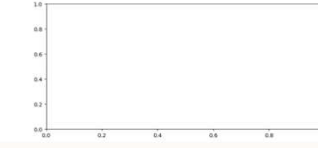
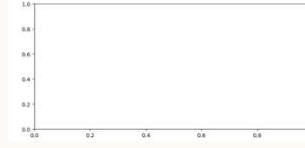
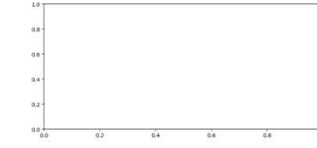
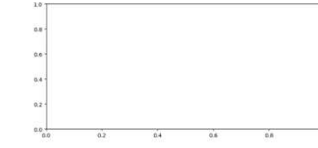
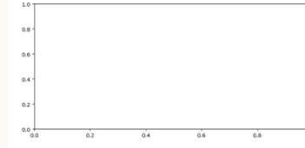
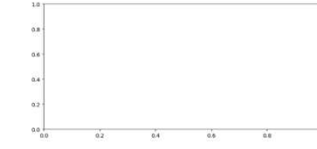
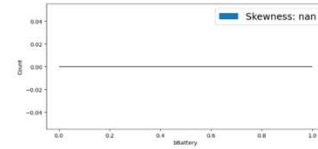
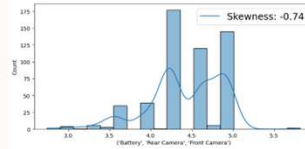
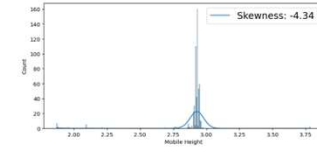
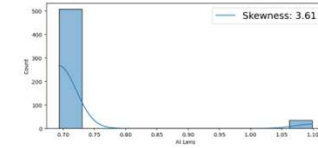
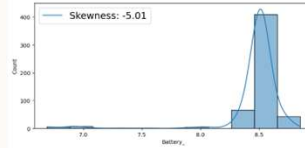
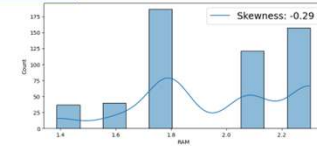
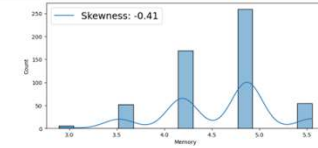
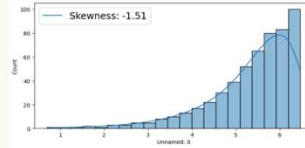
Correlation Analysis: Evaluating the correlation between different numerical features and the target variable to identify potential predictors.

Category Frequency: Analyzing the frequency of different categories in categorical features like Model, Color, and Processor to understand their prevalence in the dataset.

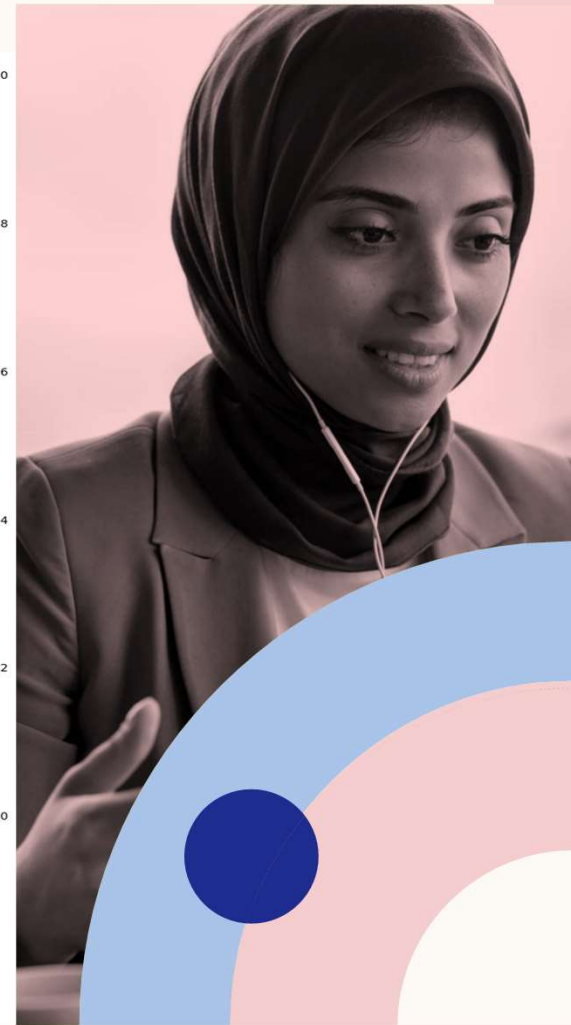
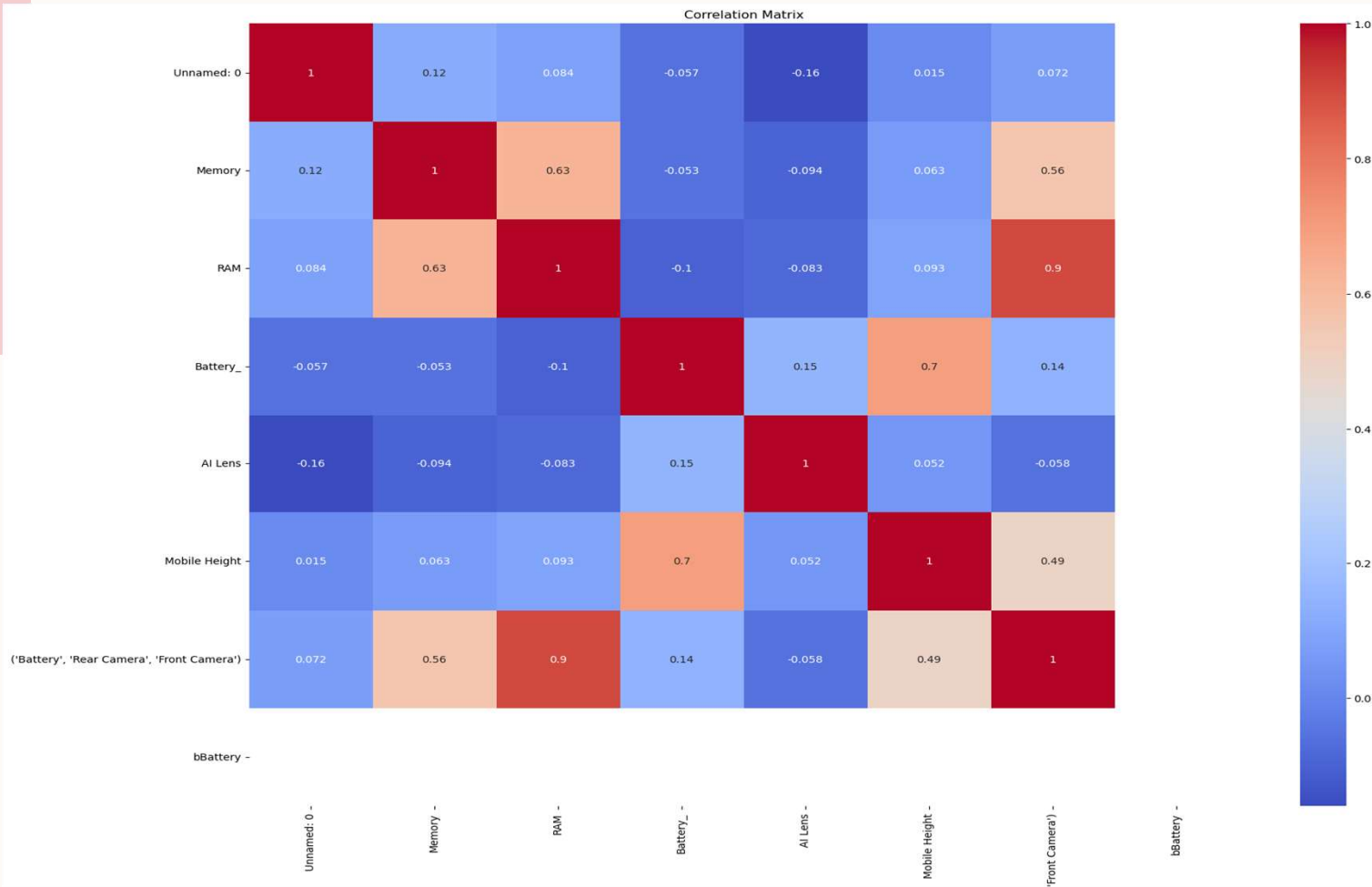


VISUALIZATION





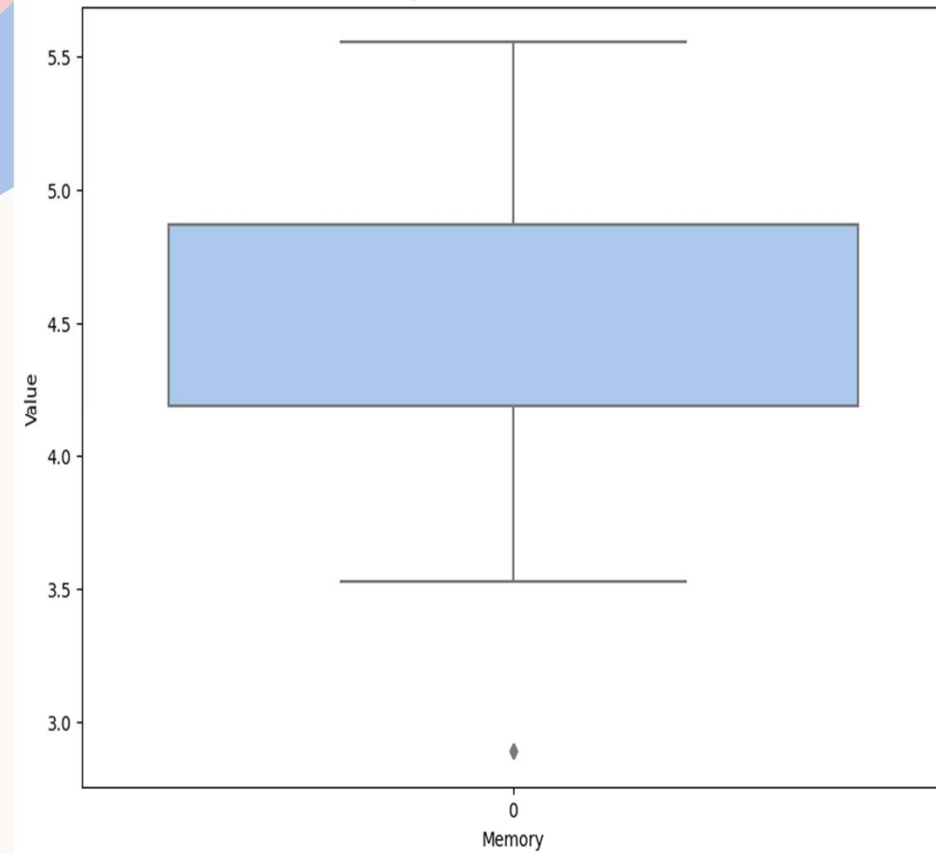
HEAT MAP



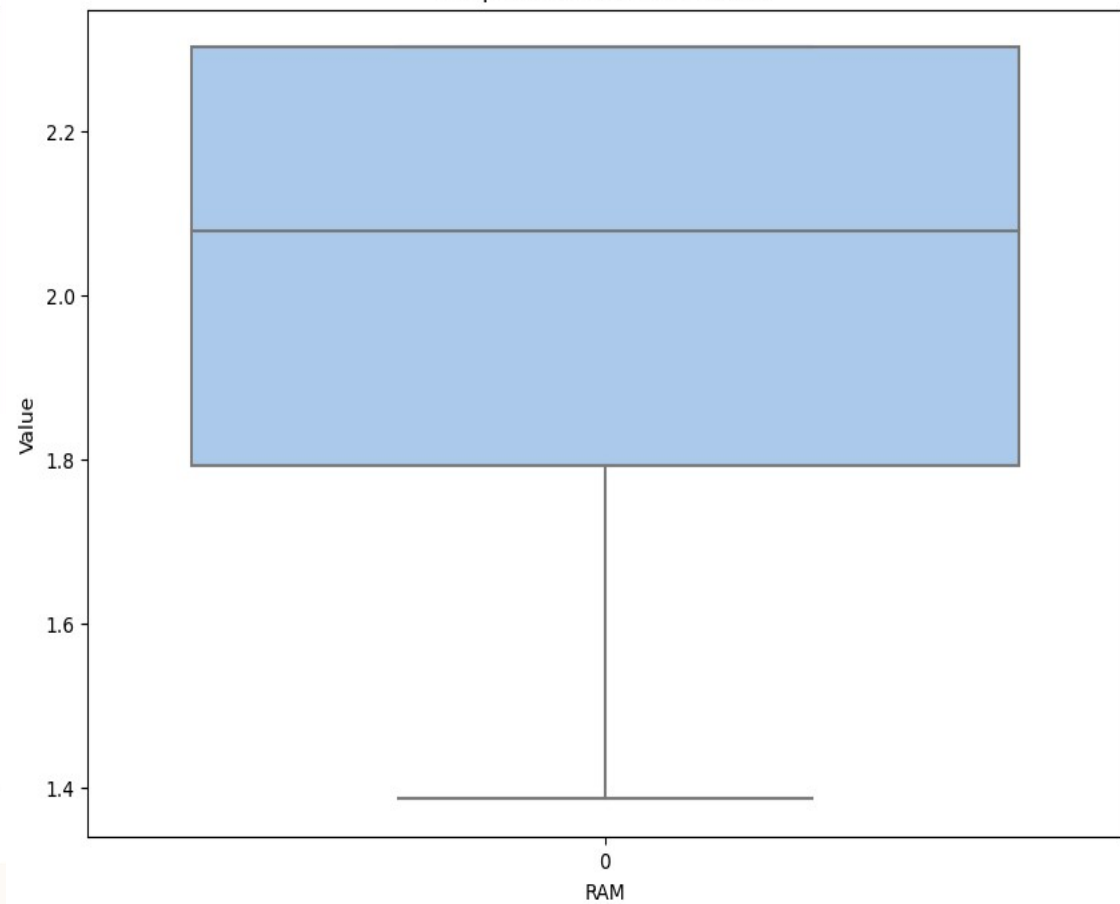
BOXPLOT

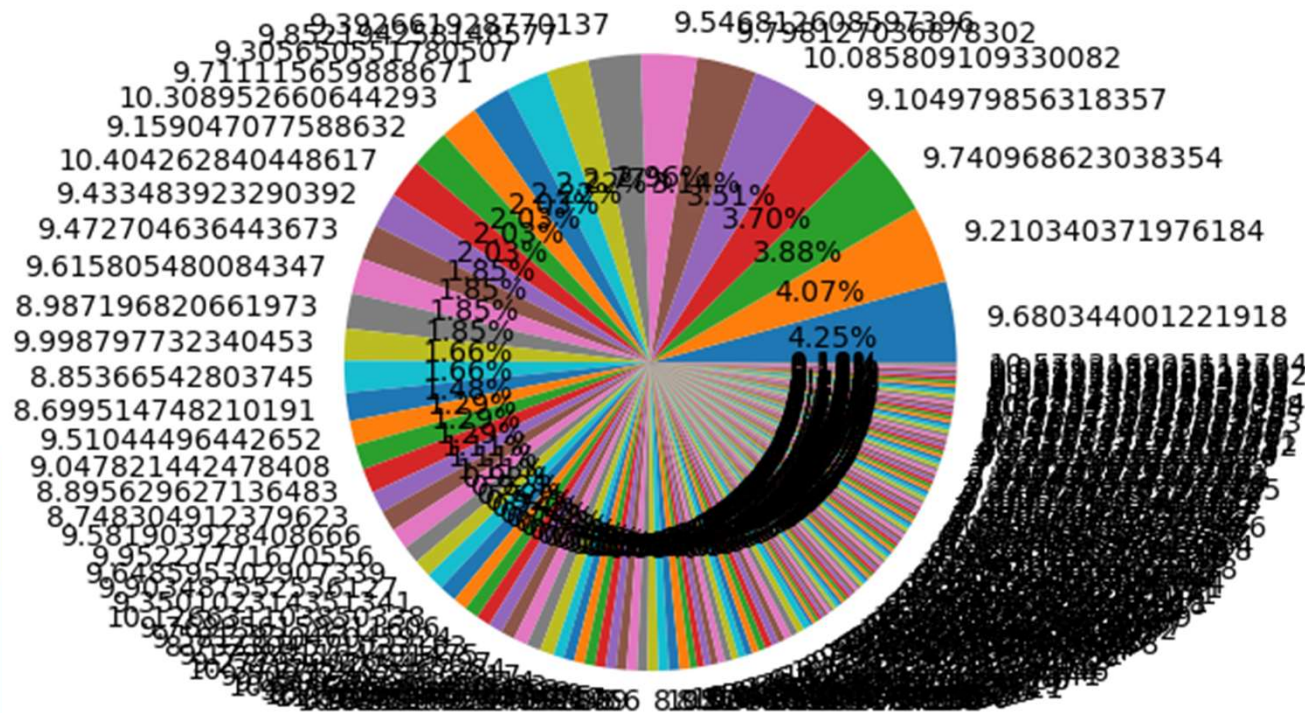
9

Boxplots of Numeric Columns



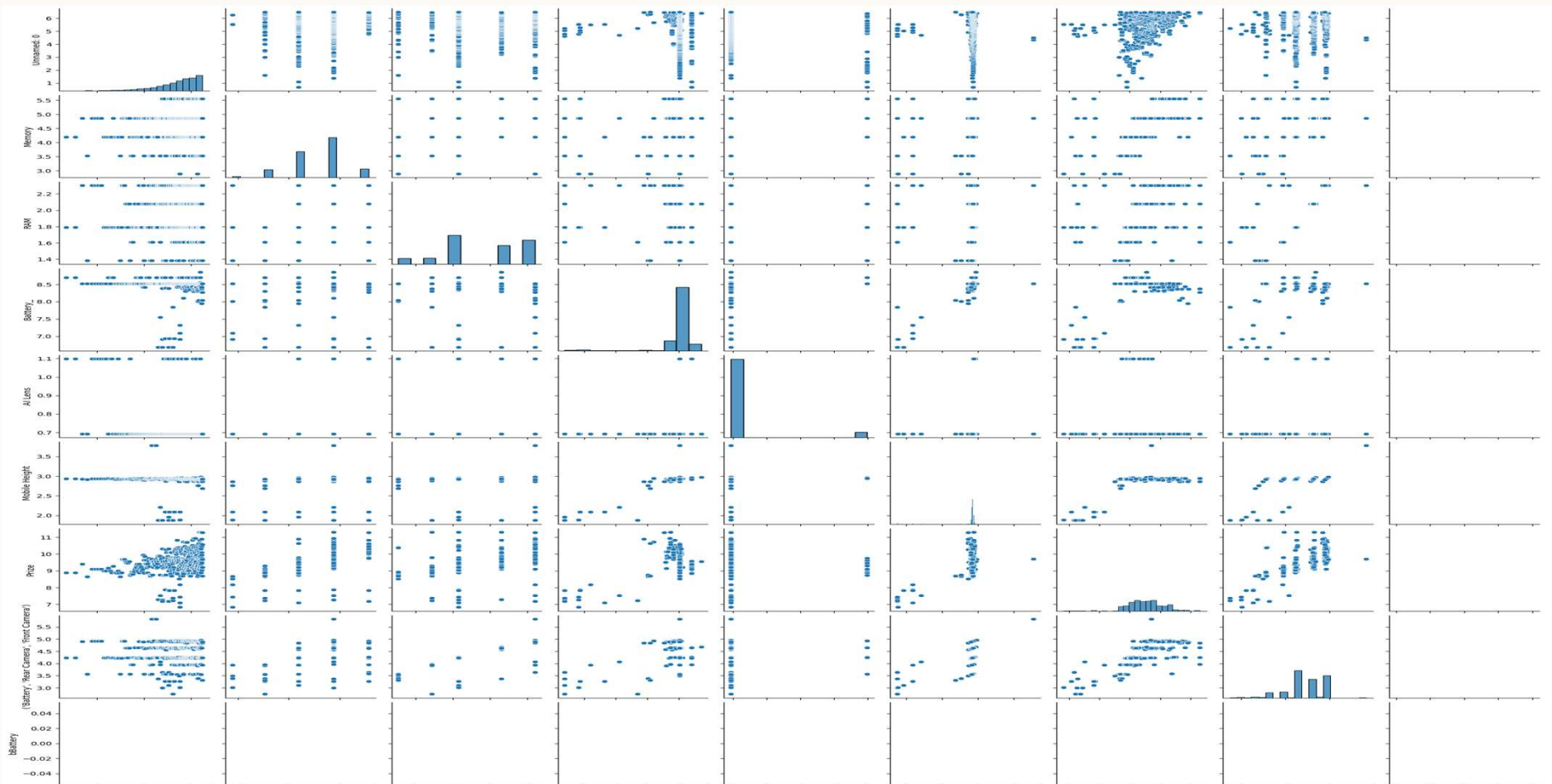
Boxplots of Numeric Columns





PAIR PLOT

11



The background features a large, light cream-colored circle on the left and a large, light pink circle on the right. These two circles overlap in the center. The area where they overlap is filled with a dark blue color. In the top right corner, there are several concentric white circles of varying radii, creating a ripple effect. The overall design is minimalist and modern.

THANK YOU

Nisha Pal