ReCell Project
Business Presentation

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

Contents

- Business Problem Overview
- Data Overview
- Exploratory Data Analysis (EDA)
- Model Performance Summary
- Business Insights and Recommendations

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

Business Problem Overview and Solution Approach

- ReCell is a startup aiming to tap the rising potential of the comparatively under-the-radar used phone and tablet market, which is predicted to be worth $52.7bn by 2023 by the International Data Corporation.
- Refurbished and used devices provide cost-effective alternatives to both consumers and businesses that are looking to save money when purchasing a phone or tablet. They can be sold with warranties and can also be insured with proof of purchase. Maximizing the longevity of devices through second-hand trade also reduces their environmental impact and helps in recycling and reducing waste.
- With third-party vendors/platforms, such as Verizon, Amazon, etc., providing attractive offers for refurbished phones and tablets, along with the impact of the COVID-19 outbreak, the used and refurbished device market has a good chance of getting a significant boost.
- The task at hand is to analyze the data provided and develop a dynamic pricing strategy for used and refurbished phones and tablets using a linear regression model and identify factors that significantly influence the price of a used device.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

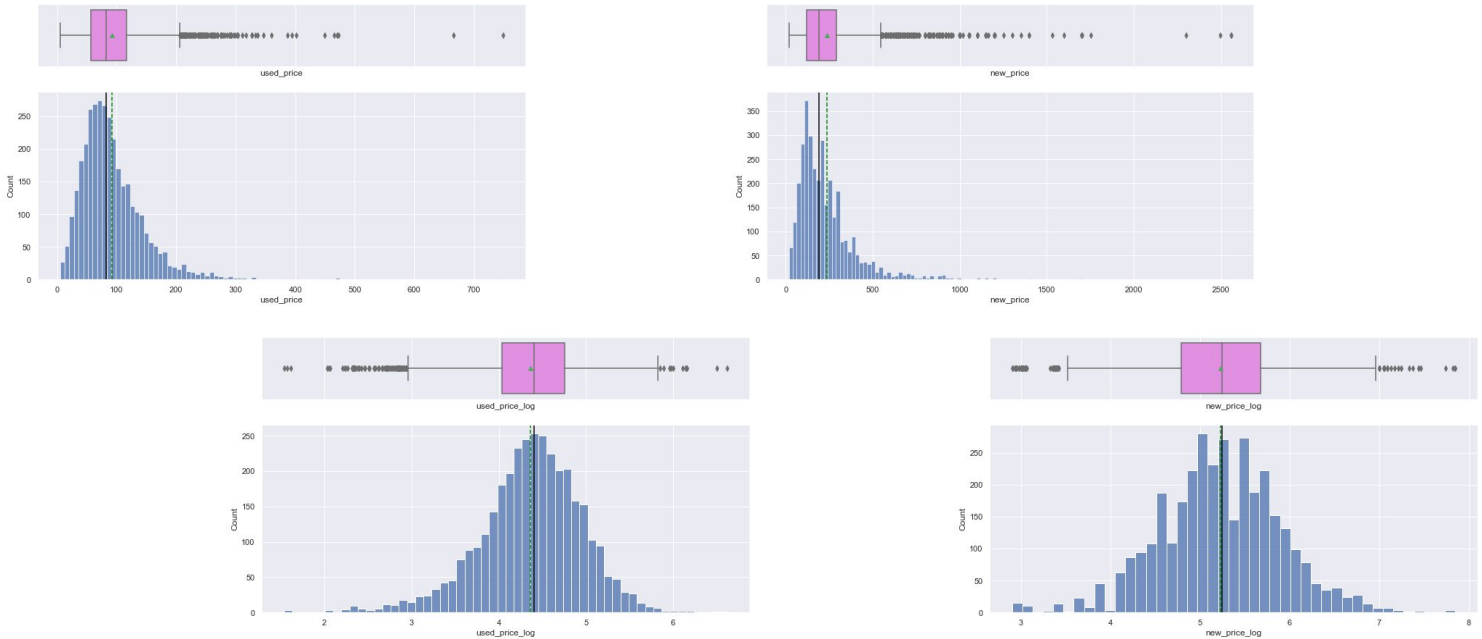
Data Overview

- The data contains information about 3571 phones and tablets and their characteristics.
- The characteristics include brand name, availability of 4G and 5G, screen size, weight, storage capacity, amount of RAM, used model price, new model price, and more.
- We will determine price segments based on the price of a new model of the used device being sold.
- There are a few missing values in some numerical columns. We will be imputing them as follows:
    - Initially, missing values will be imputed by the median grouped by release year and brand.
    - The remaining missing values will be imputed by the median grouped by brand.
    - Any leftover missing values will be imputed by the column median.
- We will also apply a log transform to some of the variables to deal with skewness in them.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

Exploratory Data Analysis

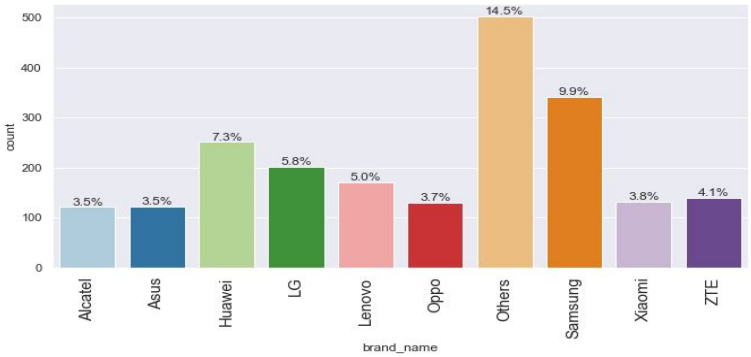
- The prices of used devices and their new models are heavily skewed.
- Log transform has been applied to both to reduce the extreme skewness.

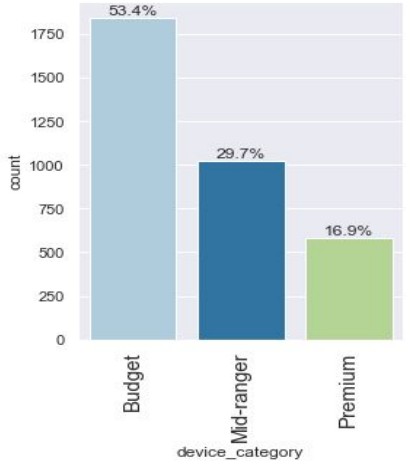


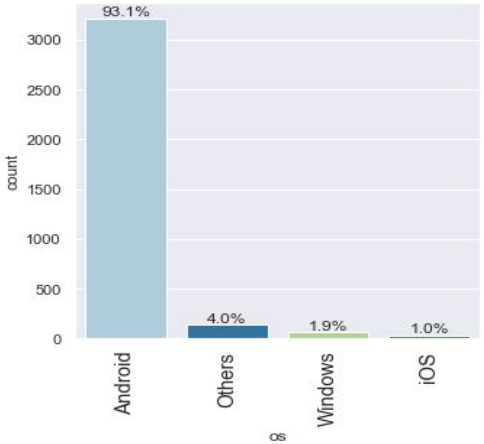
Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

Exploratory Data Analysis

- Samsung has the most number of devices in the data, followed by Huawei and LG.
- Android phones and tablets dominate more than 90% of the market.
- More than 50% of the devices in the data are budget devices.



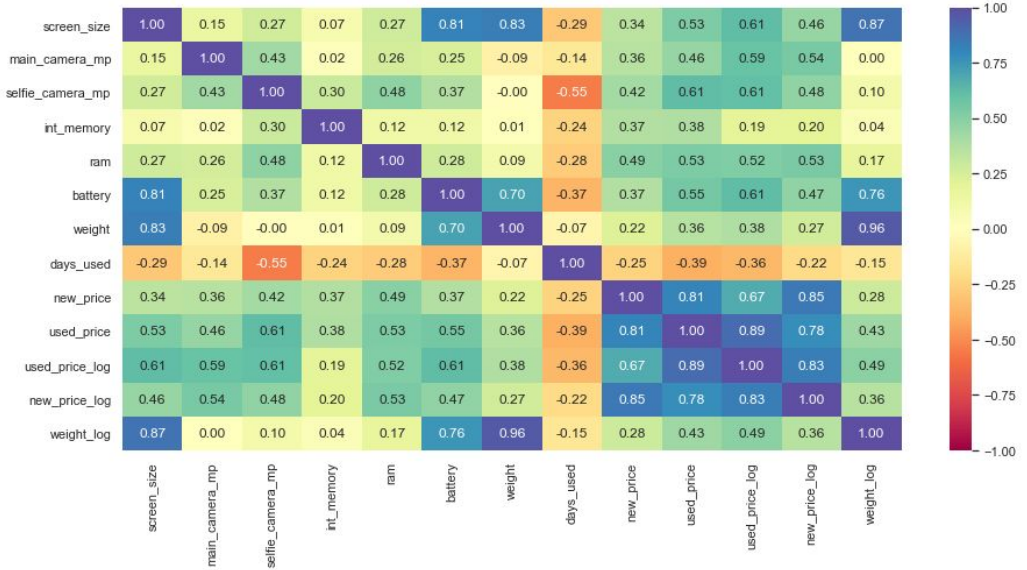




Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

Exploratory Data Analysis

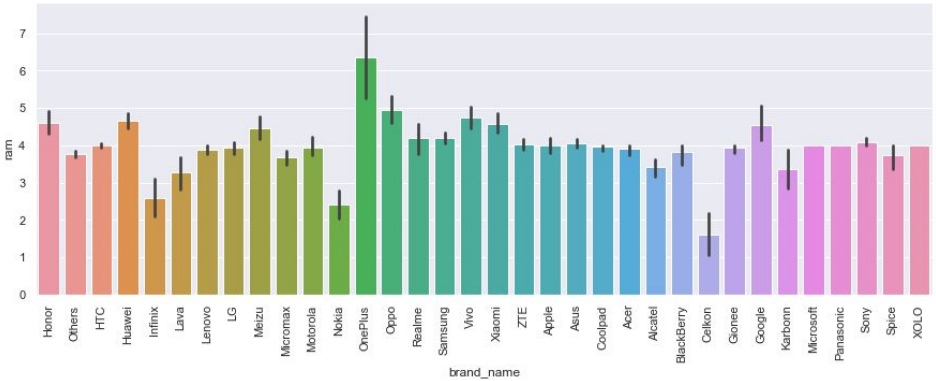
- The used device price is highly correlated with the price of a new device model. This makes sense as the price of a new model is likely to affect the used device price.
- Weight, screen size, and battery capacity of a device show a good amount of correlation. This makes sense as larger battery capacity requires bigger space, thereby increasing screen size and weight.
- The number of days a device is used is negatively correlated with the resolution of its front camera. This makes sense as older devices did not offer as powerful front cameras as the recent ones.



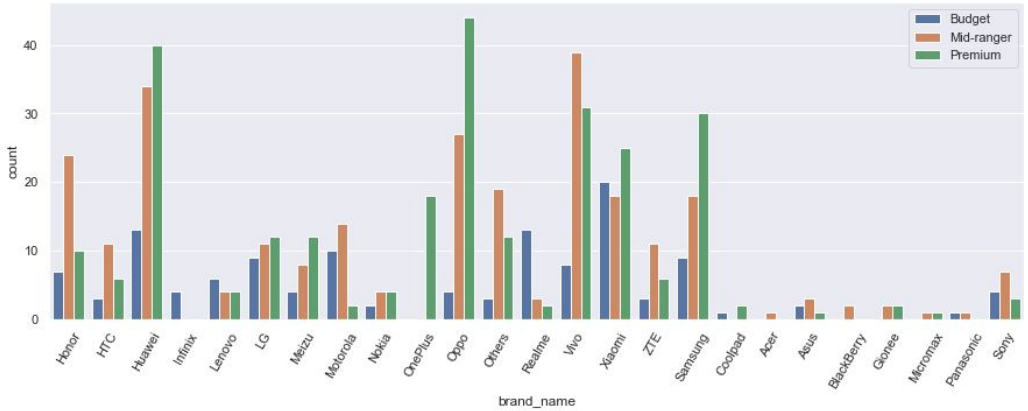
Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

Exploratory Data Analysis

- Most of the companies offer around 4GB of RAM on average. OnePlus offers the highest amount of RAM in general, while Celkon offers the least.



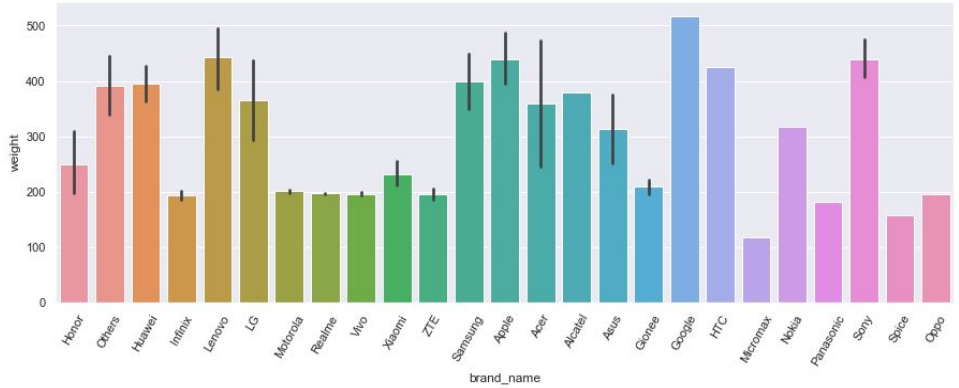
- Customers specifically look for good front cameras to click cool selfies can consider Huawei as the go-to brand as they offer many phones across different price ranges with powerful front cameras.



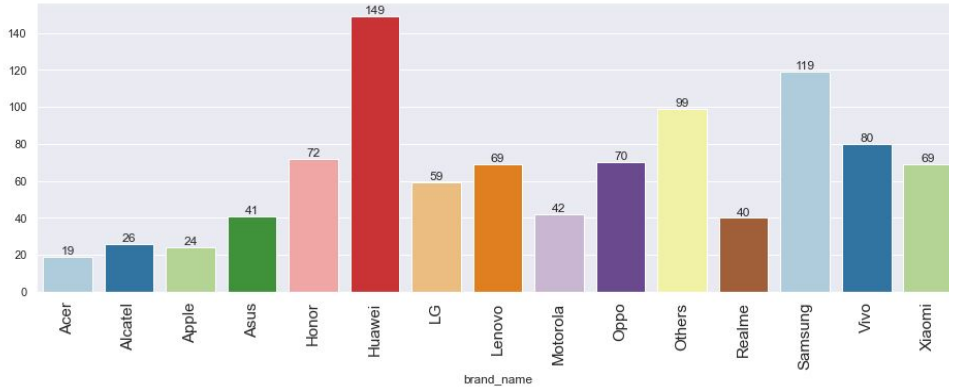
Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

Exploratory Data Analysis

- A lot of brands offer devices which are not very heavy but have a large battery capacity. Some devices offered by brands like Vivo, Realme, Motorola, etc. weigh just about 200g but offer great batteries. Some devices offered by brands like Huawei, Apple, Sony, etc. offer great batteries but are heavy.



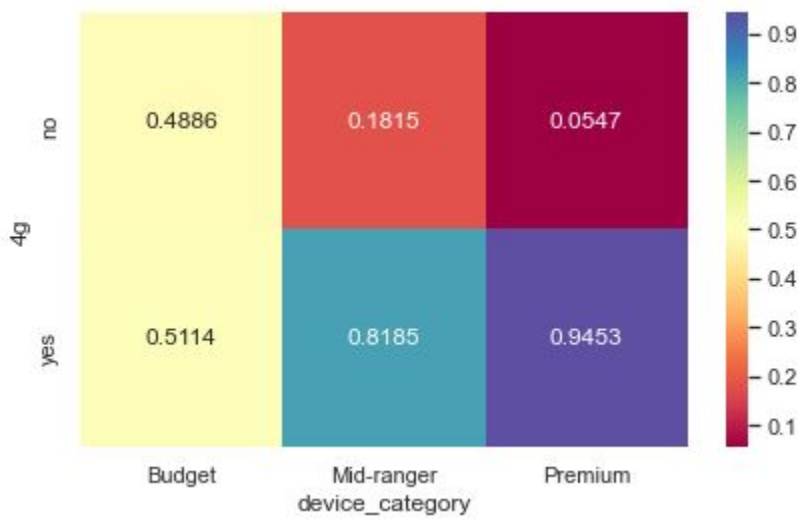
- Huawei and Samsung offer a lot of devices suitable for customers buying phones and tablets with bigger screens for entertainment purposes.

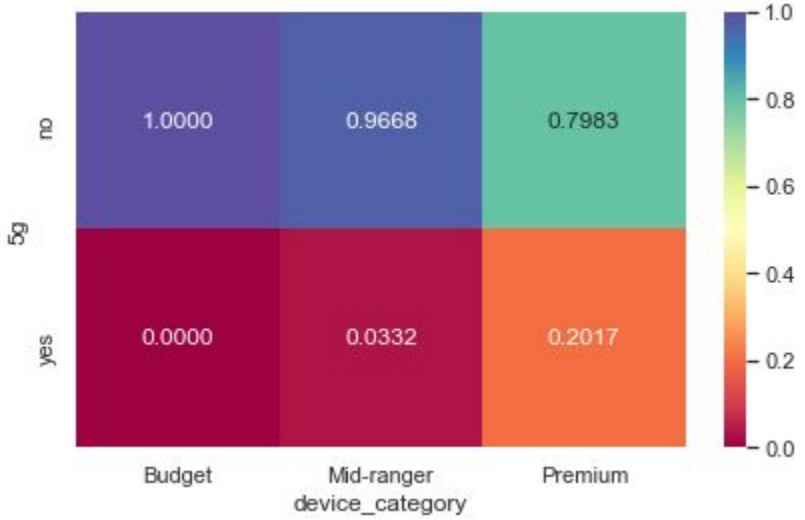


Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

Exploratory Data Analysis

- There is an almost equal number of 4G and non-4G budget devices, but there are no budget devices offering 5G network.
- Most of the mid-rangers and premium devices offer 4G network.
- Very few mid-rangers (~3%) and around 20% of the premium devices offer 5G network.





Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

Model Performance Summary

- We want to predict the price of a used device based on the characteristics provided to us.
- We have used the normalized version 'used_price_log' for building a robust Linear Regression model using the train data and check the performance on test data to understand the predictive power of our model.
- The model satisfies all the assumptions of Linear Regression and indicates that the most significant predictors of the used phone price are:
    - Price of a new phone of the same model
    - Device screen size
    - Front and rear camera resolutions
    - Amount of RAM
    - Number of days it was used
    - Availability of 4G and 5G network

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

Model Performance Summary

- We have got an R-squared and adjusted R-squared of ~0.8, which is a clear indication that we have been able to create a very good model that is able to explain variance in the price of used phones up to 80%.

Data	RMSE	MAE	MAPE
Train	40.90	18.93	20.90
Test	26.92	17.57	20.96

- Mean Absolute Error indicates that our current model is able to predict used phone prices within a mean error of 17.5 euros on the test data.
- MAPE is around 21 on the test data, which means that we are able to predict within 21% of the price value.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

Business Insights and Recommendations

- The model explains 80% of the variation in the data and can predict within 17.5 euros of the used phone price.
- The most significant predictors of the used device price are the price of a new device of the same model, the size of the devices screen, the resolution of the rear and front cameras, the number of days it was used, the amount of RAM, and the availability of 4G and 5G network.
- A unit increase in new model price will result in a 0.09% increase in the used device price.
- A unit increase in size of the device's screen will result in a 5.76% increase in the used device price.
- A unit increase in the amount of RAM will result in a 3.69% increase in the used device price

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

Business Insights and Recommendations

- ReCell should look to attract people who want to sell used phones and tablets which have not been used for many days and have good front and rear camera resolutions.
- Devices with larger screens and more RAM are also good candidates for reselling to certain customer segments.
- They should also try to gather and put up phones having a high price for new models to try and increase revenue.
    - They can focus on volume for the budget phones and offer discounts during festive sales on premium phones.
- Additional data regarding customer demographics (age, gender, income, etc.) can be collected and analyzed to gain better insights into the preferences of customers across different segments.
- ReCell can also look to sell other used gadgets, like smart watches, which might attract certain segments of customers.

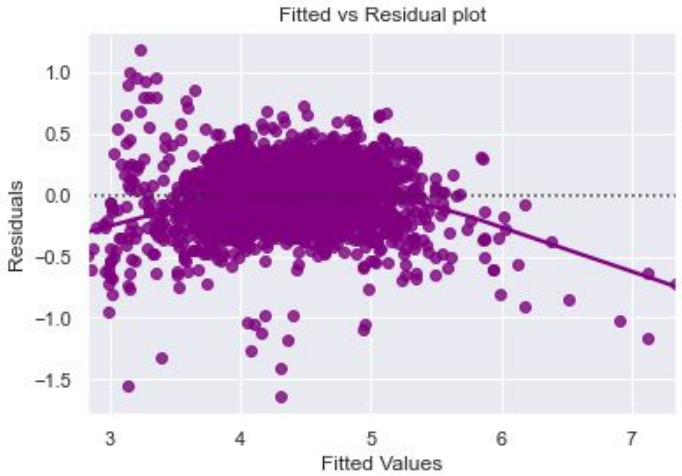
Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

greatlearning
Power Ahead

Happy Learning !

Appendix: Model Assumptions

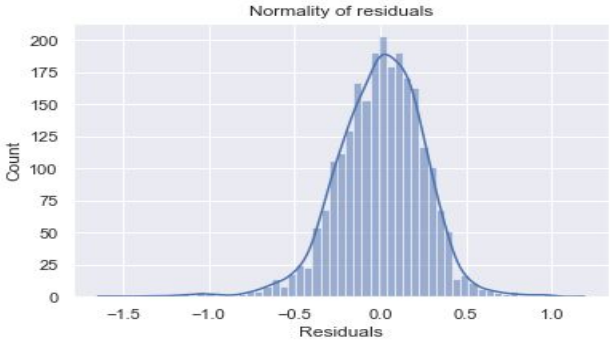
- We have checked for multicollinearity using VIF.
  - All numerical variables with VIF > 5 were dropped.
  - VIF for dummy variables were ignored.
- For linearity and independence tests, we have checked the residuals vs fitted values plot.
  - The plot does not show any pattern, so the assumption is satisfied.

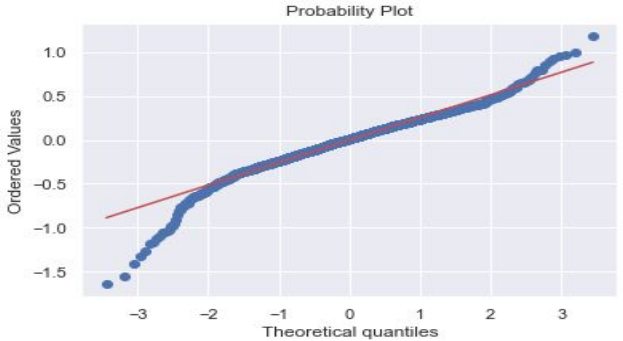


Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

Appendix: Model Assumptions

- For normality test, we checked the distribution of residuals, the Q-Q plot, and also applied the Shapiro-Wilk test
    - The distribution does seem close to normal as per the plots.
    - The Shapiro-Wilk test does suggest that the distribution is not strictly normal, but we will accept this an approximately normal distribution.





- For homoscedasticity test, we applied the goldfeldquadt test, which showed that the residuals were homoscedastic.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.