

# ReCell – Supervised Learning Business Presentation

# Contents

The presentation consists of three Four Sections:

1. Background and Business Problem Overview
2. Data Overview
3. Exploratory Data Analysis
4. Model overview and performance summary
5. Key findings and insights

# Background of the Business and Objective

## Background

Used and refurbished phone market has grown considerably over the past decade.

Third-party vendors/platforms, such as Verizon, Amazon, etc., provide attractive offers to customers for refurbished smartphones.

The impact of the COVID-19 outbreak may further boost the cheaper refurbished smartphone segment, as consumers cut back on discretionary spending and buy phones only for immediate needs.

**ReCell, a startup aiming to tap the potential in this market.**

## Objective

1. Do statistical analysis and extract actionable insights from the data
2. Analyze the data and build a linear regression model to predict the price of a used phone and identify factors that can significantly influence it.
3. To come up with a ML-based solution to develop a dynamic pricing strategy for used and refurbished smartphones.

# Data Overview

Data	Description
brand_name	Name of manufacturing brand
os	OS on which the phone runs
screen_size	Size of the screen in cm
4g	Whether 4G is available or not
5g	Whether 5G is available or not
main_camera_mp	Resolution of the rear camera in megapixels
selfie_camera_mp	Resolution of the front camera in megapixels
int_memory	Amount of internal memory (ROM) in GB
Ram	Amount of RAM in GB
battery	Energy capacity of the phone battery in mAh
weight	Weight of the phone in grams
release_year	Year when the phone model was released
days_used	Number of days the used/refurbished phone has been used
new_price	Price of a new phone of the same model in euros
used_price	Price of the used/refurbished phone in euros

Observation	Variable
3571	15

## Note:

1. Missing or null values in below columns

main\_camera\_mp

selfie\_camera\_mp

int\_memory

ram

battery

Weight

2. Missing value treatment has been done for main\_camera\_mp.

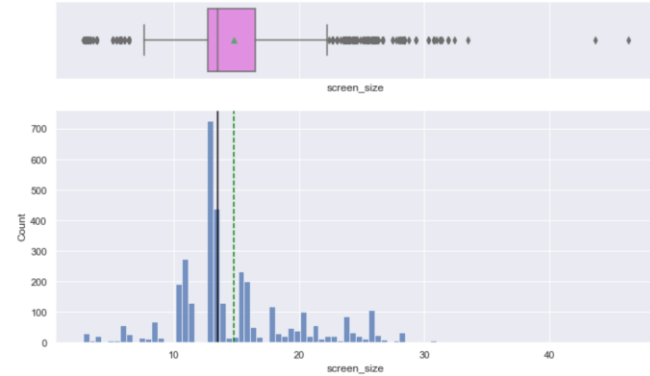
3. brand\_name, os, 4g, 5g are objects, these are converted to categories.

# Exploratory Data Analysis

- There are 33 different brand names which comes into 125 screen sizes.
- Screen size varies from 2.7 inches to 46.36 CMS.
- All 33 brands released in last 8 different years.
- Used price also varies between 2.51 to 1916.54 Euros.

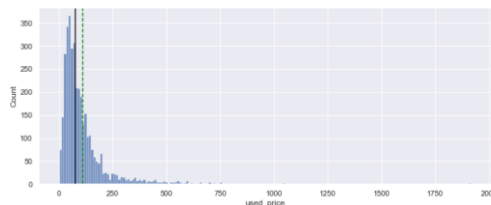
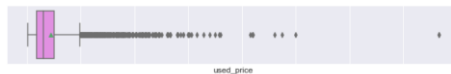
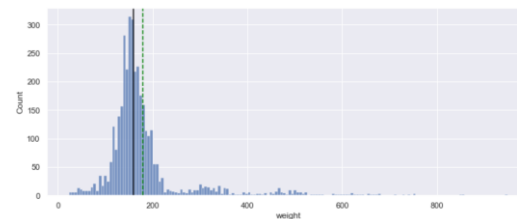
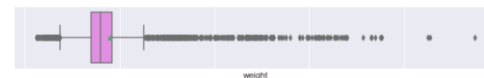
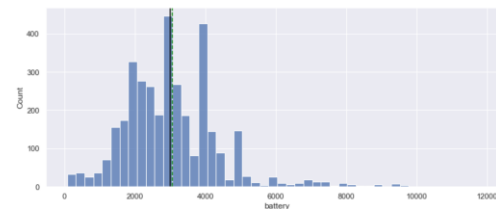
# Univariate Analysis –Screen Size

- Highest screen size for used smart phone are in the range of 10.16 to 13.49 cms.
- There are outliers in the screen size.
- Used phones with screen size 13.49 cms has the highest count.



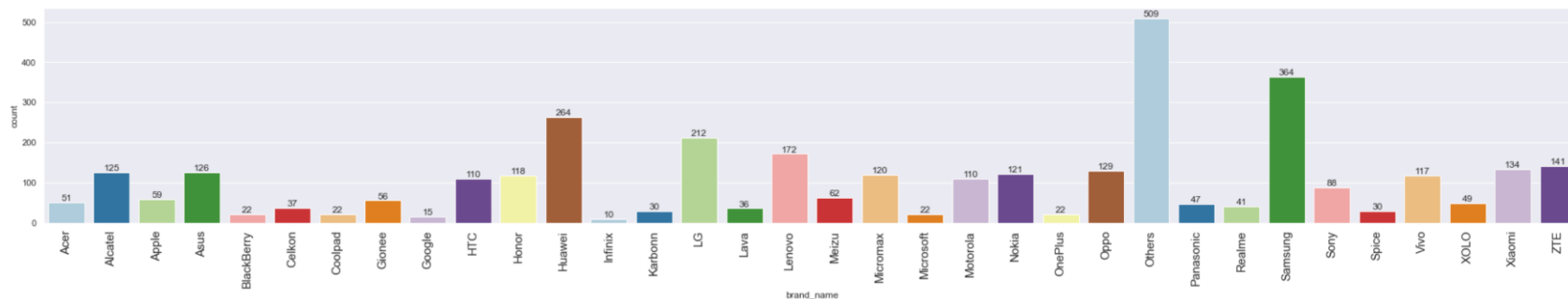
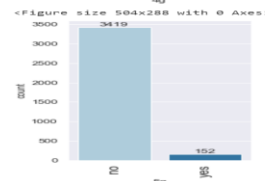
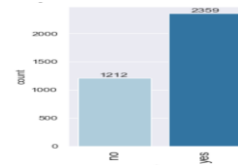
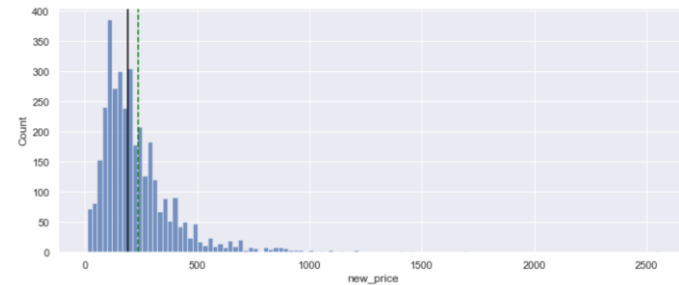
# Univariate Analysis

- There are outliers exists with the battery
- Maximum used phones have a battery capacity ranged between 2000-4000 mAh.
- There is no skewness in battery.
- Used phones weight also has a lot of outliers but the average weight is 179 grams.
- Used phones used\_price has many outliers and it is right skewed.
- Used phones new\_price has many outlier and it is also right skewed.



# Univariate Analysis

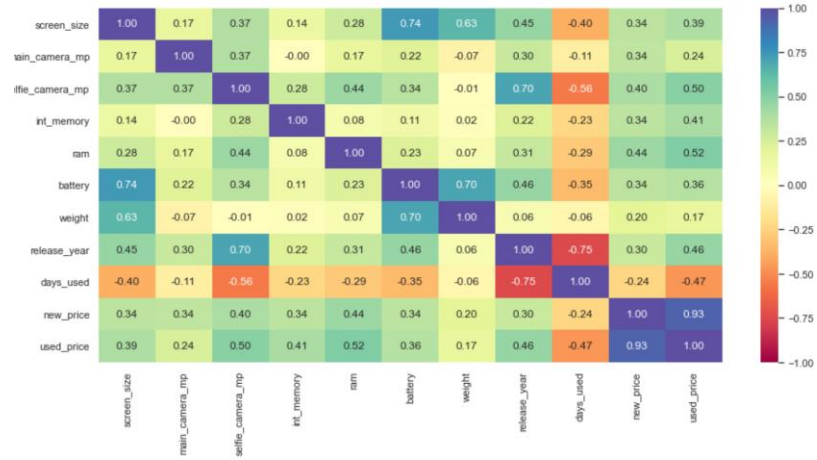
- Used phones new\_price has many outlier and it is also right skewed.
- Days used for used phones is left skewed.
- New price for used phones is also rightly skewed.
- Used price is also rightly skewed
- Huawei, LG Samsung and Others are the highest available brands of cell phones.
- Used phones are most popular with android OS.
- Used phones mostly available in 4G spectrum.
- Used phones with 5G availability is very less.





# Bivariate Analysis

- Screen Size with battery are positively correlated.
- New price is highly positively correlated with used price.



# Model overview and performance summary

- We used regular Linear regression and Linear Regression using statsmodels to build our machine learning model.
- The training  $R^2$  is 96%, indicating that the model explains 96% of the variation in the train data. So, the model is not underfitting.
- MAE and RMSE on the train and test sets are comparable, which shows that the model is not overfitting.
- MAE indicates that our current model is able to predict used\_price within a mean error of 10.41 Euros on the test data.
- MAPE on the test set suggests we can predict within 18.1% of the used\_price.

# Model overview and performance summary

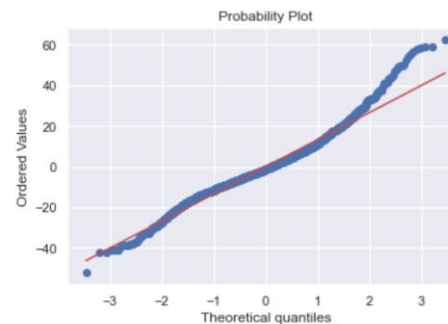
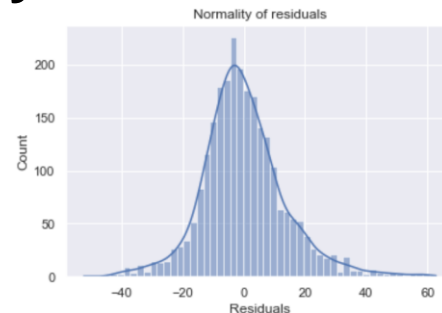
- brand\_name\_Apple, brand\_name\_Others, brand\_name\_Samsung and os\_iOS have VIF greater than or equal to 5.
- brand\_name\_Apple and os\_iOS have a VIF score of much greater than 10 i.e. 23 and 22 which shows signs of high multicollinearity.
- This does seem to make intuitive sense because the Apple brand and OS IOS are highly co-related.

# Model overview and performance summary

- brand\_name\_Apple, brand\_name\_Others, brand\_name\_Samsung and os\_iOS have VIF greater than or equal to 5.
- brand\_name\_Apple and os\_iOS have a VIF score of much greater than 10 i.e. 23 and 22 which shows signs of high multicollinearity.
- This does seem to make intuitive sense because the Apple brand and OS IOS are highly co-related.

# Model overview and performance summary

- The residuals were more or less followed a straight line except for the head and tails.
- The residual data followed normal distribution with the model we followed.
- We observed that our model returned reasonable prediction results, and the actual and predicted values are slightly comparable though not fully.
- The model is able to explain ~95.3% of the variation in the data, which is very good.



# Key Findings and Insights

Used price is heavily dependent on New Price, RAM, Internal Memory, Selfie Camera Megapixel. As these increase, the total used price increases, as is visible in the positive coefficient sign.

The used price is negatively correlated with Infinix, Gionee, Lenovo, OnePlus brands, Which means that for these brands used price decreases.

Brand\_name Google, Blackberry and LG are positively impacting the used phone pricing.  
Phones with IOS operating systems (Apple) have highest used price in the market.

Screen Sizes have positive influence on used price with increase in screen size there is an increase in used price.

**greatlearning**  
*Power Ahead*

Happy Learning !

