ReCell SLF Project Business Presentation

Contents

Business Problem Overview and Solution Approach

- Core Business Idea - Buying and selling used smartphones used to be something that happened on a handful of online marketplace sites. But the used and refurbished phone market has grown considerably over the past decade, and a new IDC (International Data Corporation) forecast predicts that the used phone market would be worth $52.7bn by 2023 with a compound annual growth rate (CAGR) of 13.6% from 2018 to 2023. Refurbished and used devices continue to provide cost-effective alternatives to both consumers and businesses that are looking to save money when purchasing a smartphone. There are plenty of other benefits associated with the used smartphone market. Used and refurbished devices can be sold with warranties and can also be insured with proof of purchase.
- Problem to tackle - The rising potential of refurbished smartphone segment comparatively under-the-radar market fuels the need for an ML-based solution to develop a dynamic pricing strategy for used and refurbished smartphones. ReCell, a startup aiming to tap the potential in this market, wants to analyze the data provided and build a linear regression model to predict the price of a used phone and identify factors that significantly influence it.

Data Overview

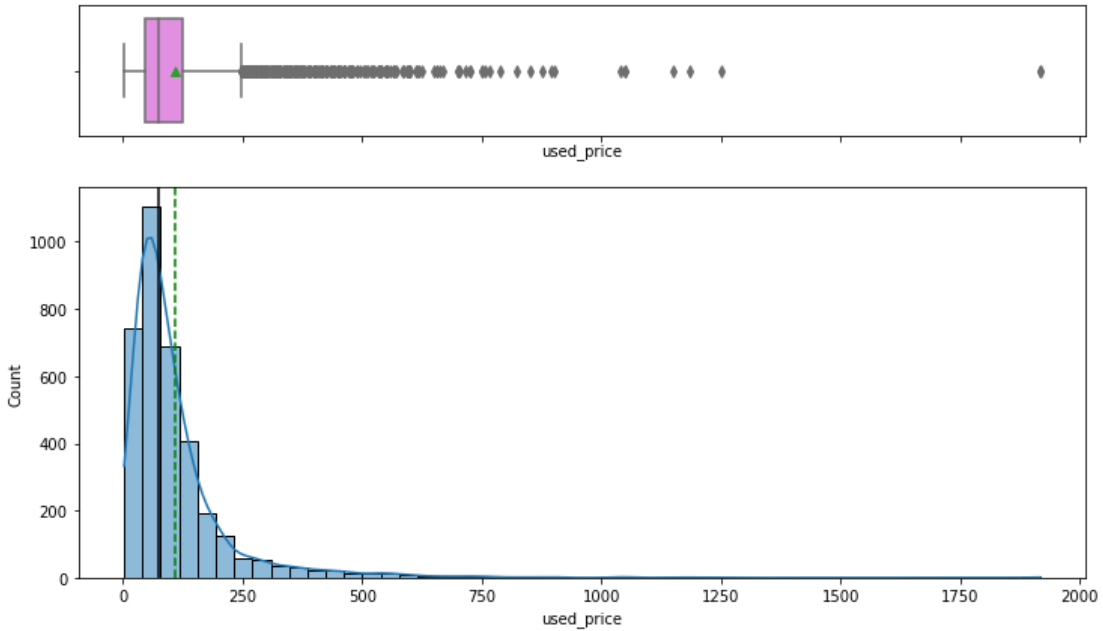
- The data contains the different attributes of used/refurbished phones.

Data Dictionary
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

Exploratory Data Analysis (EDA)

- Univariate Analysis - Dependent Variable – Used Price

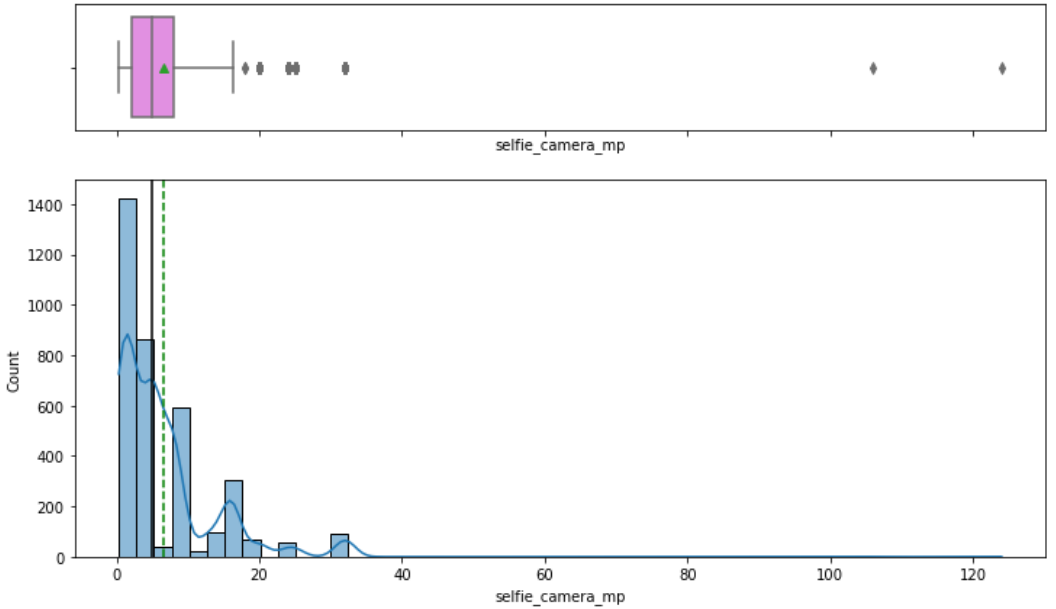
  - There are plenty of outliers in the data as displayed in box plot.
  - The mean is greater than median, that indicates the distribution is skewed to the right.
  - About 68% of the values for used price are less than 125 euros.



Exploratory Data Analysis (EDA)

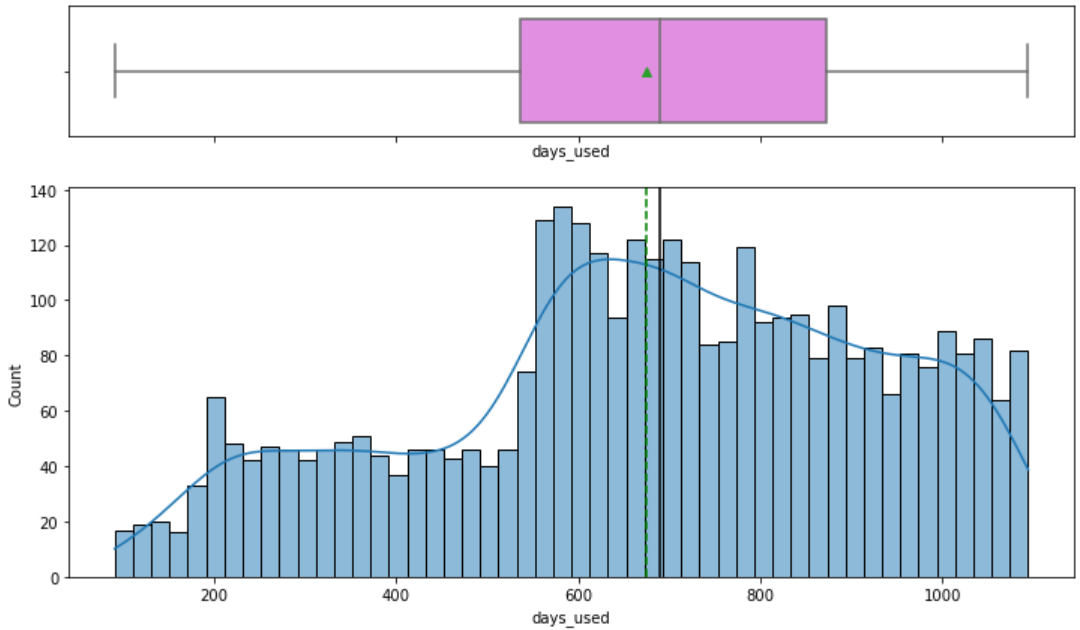
- Univariate Analysis - Independent Variable – Selfie Camera Megapixels

- There are outliers in the data as displayed in box plot.
- The mean is greater than median, that indicates the distribution is skewed to the right.
- About 68% of the values for 'selfie_camera_mp' are from 2 to 8 megapixels.



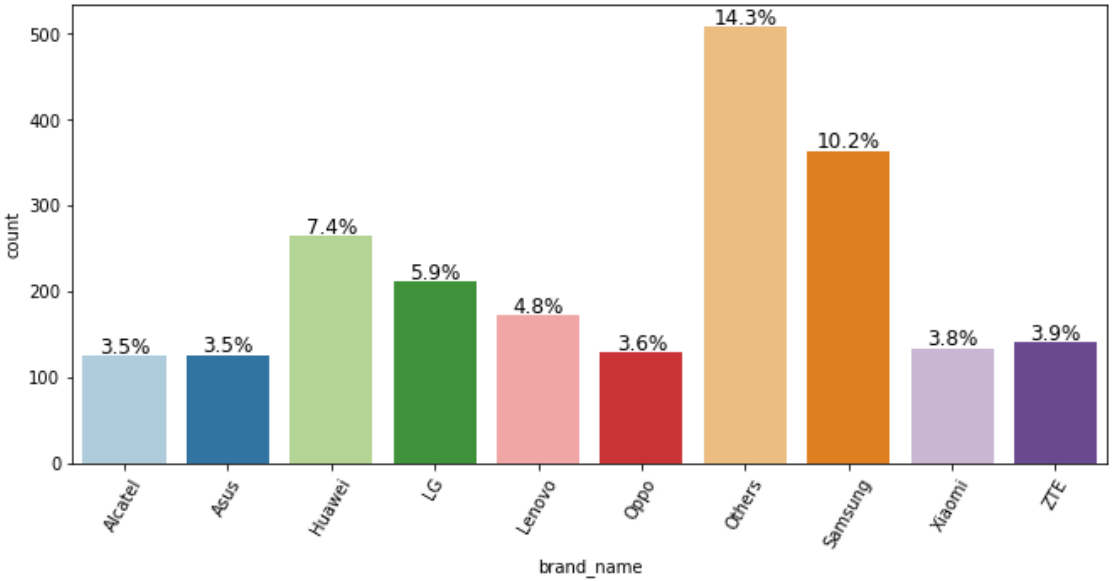
Exploratory Data Analysis (EDA)

- Univariate Analysis - Independent Variable – Days Used
  - There are no outliers in the data as displayed in box plot.
  - The mean is less than median, that indicates the distribution is skewed to the left.
  - About 68% of the values for 'days_used' are from 536 to 872 days.



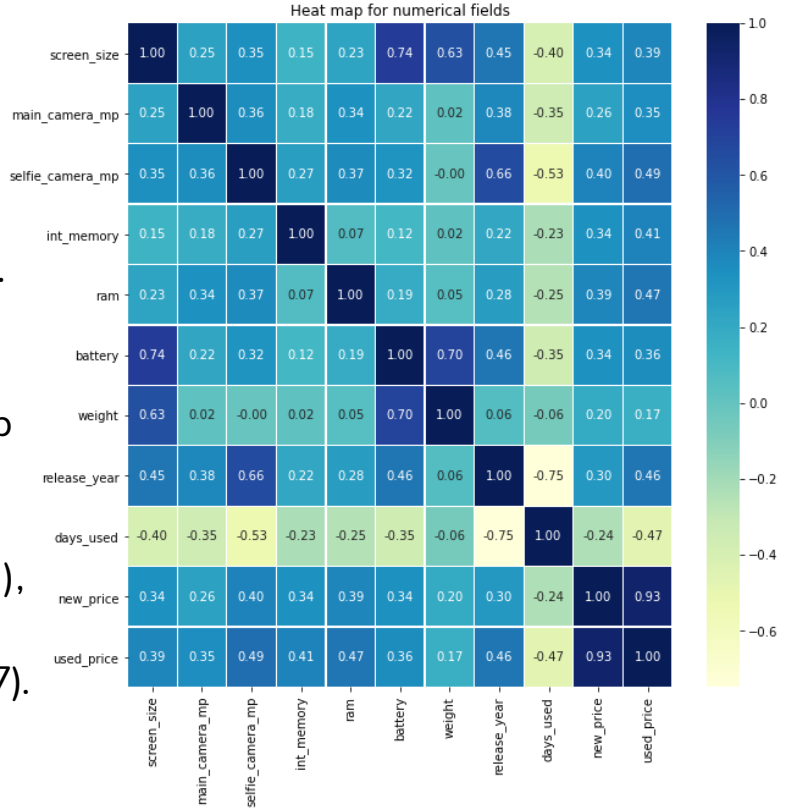
Exploratory Data Analysis (EDA)

- Univariate Analysis - Independent Variable – Brand Name
  - The field 'brand_name' has 34 unique values, one of them is "Others" that constitute to 14.3% of the data.
  - The next frequent brand is Samsung with 10.2%, followed by Huawei with 7.4%.



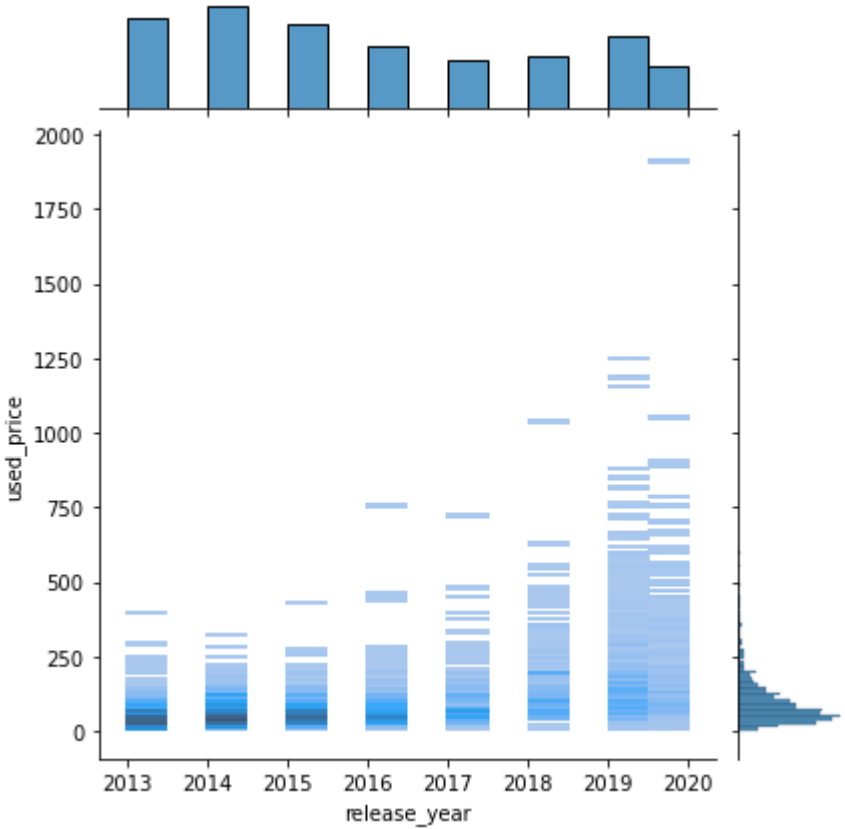
Exploratory Data Analysis (EDA) - Correlation Matrix

1. Used_price is highly positively corelated with new_price (0.93), which makes sense as the new price factors in for determining the used price of the phone.
2. Used_price is positively corelated with release_year (0.46). In other words, newer the phone, higher the price.
3. Used_price is positively corelated with ram (0.47), price increases as RAM memory is increased.
4. Used_price is positively corelated with selfie_camera_mp (0.49). In other words, higher the megapixels in selfie camera, higher the price.
5. Used_price is positively corelated with int_memory (0.41), price increases as internal memory is increased.
6. Used_price is negatively corelated with days_used (-0.47). The more the phone is used, the value goes down.



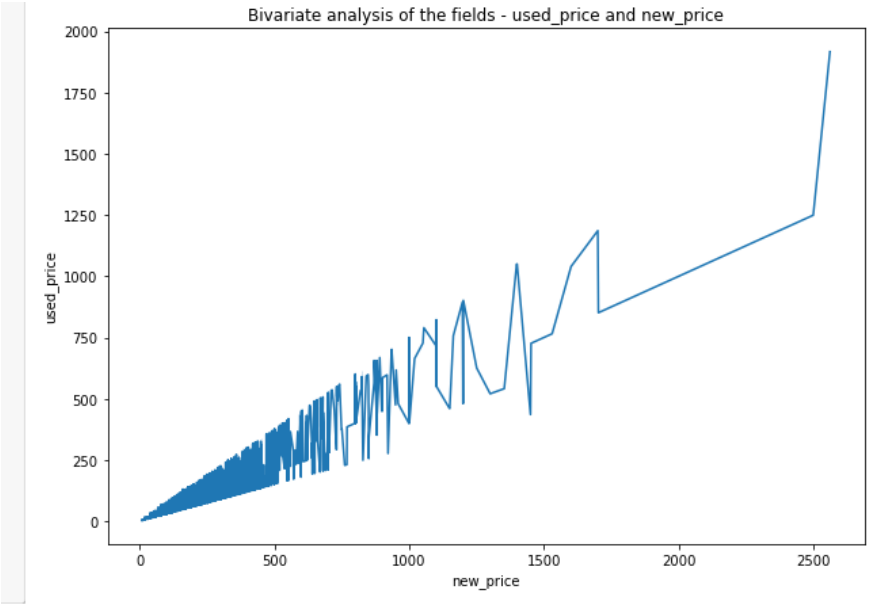
Exploratory Data Analysis (EDA)

- Bivariate Analysis – Used Price and Release Year
  - There price of the phone has increased over the years.
  - It can also be noticed that there is a higher price variation in as we move from 2013 calendar year to 2020.



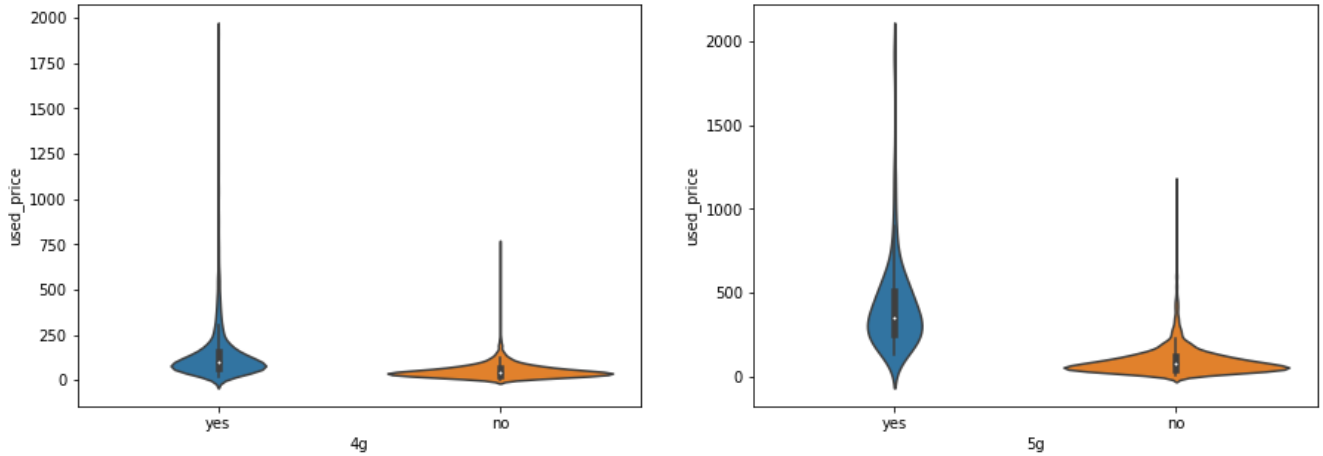
Exploratory Data Analysis (EDA)

- Bivariate Analysis – Used Price and New Price
  - As observed in the heat map, there is a sharp increase in used phone price with increase in new price.
  - Used_price is highly positively corelated with new_price (0.93), which makes sense as the new price factors in for determining the used price of the phone.



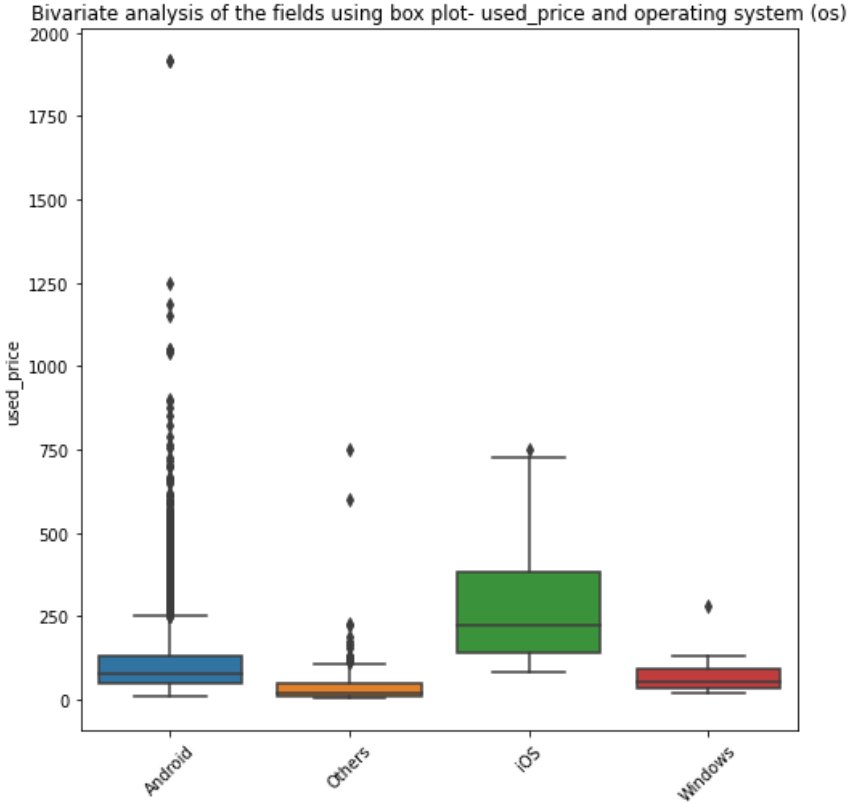
Exploratory Data Analysis (EDA)

- Bivariate Analysis – Used Price and 4G/5G
  - There is wide range of used phones with 4G feature compared to 5G which is relatively new.
  - For 4G, the price of a phone without 4G is very low compared to the phones with 4G.
  - For 5G, the price of a phone with 5G is not very high compared to a phone without 5G.



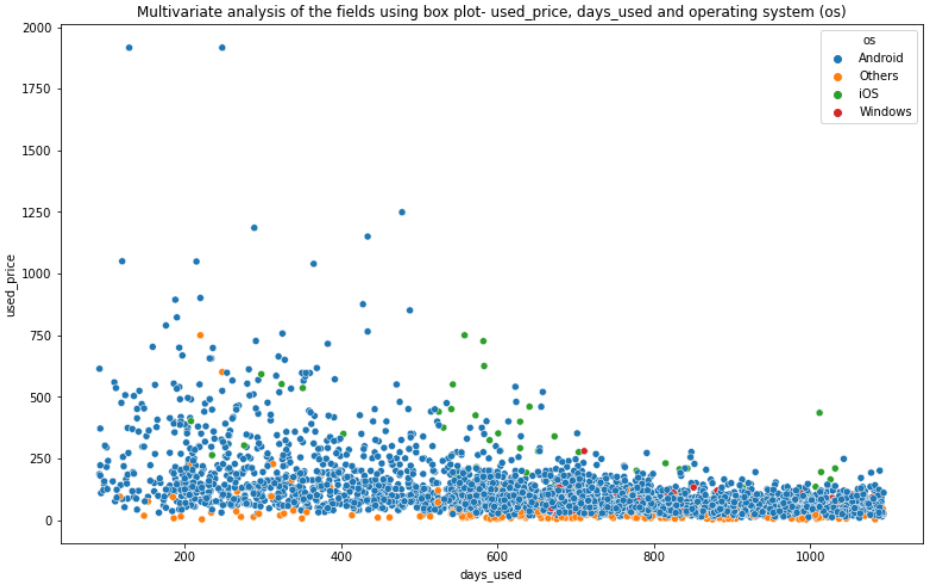
Exploratory Data Analysis (EDA)

- Bivariate Analysis – Used Price and OS
  - Phones with iOS os have higher median used prices and fewer outliers.
  - Phones with Android os have median used price in 125 euros range, but there are plenty of outliers.



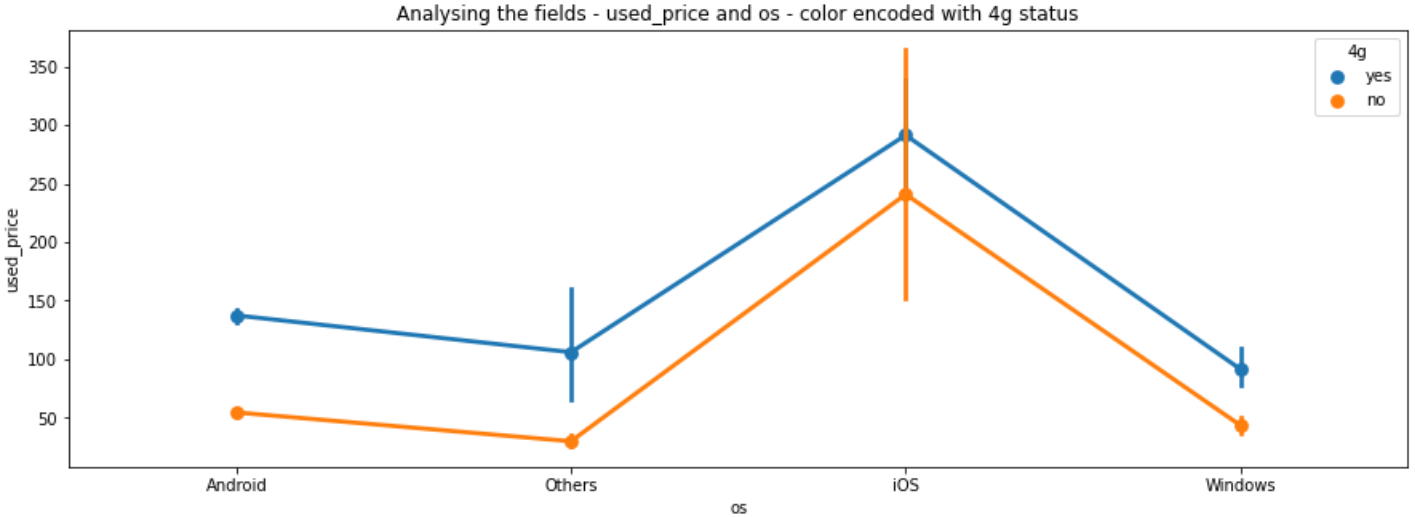
Exploratory Data Analysis (EDA)

- Multivariate Analysis – Used Price, Days Used and OS
  - It can be clearly observed that the phones with OS other than iOS, Android and Windows have lower used price for same number of days used.
  - Considering a year-old phones (360 days), iOS phones have higher price ranges compared to Android phones for same number of days used.
  - For phones less than a year old, Some of the Android phones shows higher price than iOS phones.



Exploratory Data Analysis (EDA)

- Multivariate Analysis – Used Price, 4G and OS
  - It can be clearly observed that the phones equipped with 4G have higher price across all phone operating systems.
  - Since there are a lot of observations for Android phones, we can see a large price gap between phones with and without 4g is higher. But this price gap is smaller in iOS and Windows phones.



Model Performance Summary

Model Parameters.

- The model was able to explain ~95% of the variation in the data, which is very good.
- The train and test RMSE and MAE are low and comparable. So, our model is not suffering from overfitting.
- The MAPE on the test set suggests we can predict within 17.98% of the life expectancy.

- Summary of key performance metrics for training and test data in tabular format for comparison

Training performance comparison:

	Linear Regression sklearn	Linear Regression statsmodels
RMSE	13.749809	13.841573
MAE	10.186292	10.230734
R-squared	0.956775	0.956196
Adj. R-squared	0.955946	0.955878
MAPE	18.037741	18.120497

Test performance comparison:

	Linear Regression sklearn	Linear Regression statsmodels
RMSE	14.539187	14.276996
MAE	10.530216	10.340656
R-squared	0.951679	0.953406
Adj. R-squared	0.949461	0.952609
MAPE	18.574210	17.981710

Model Performance Summary

Summary of most important factors used by the ML model for prediction

1. No of days used has a negative correlation to unit price; the less days the phone is used, higher the price tag on it.
2. The 5G phones have positive correlation and 4G phones have negative correlation. As 5G is an emerging feature, the phones with 5G capability can be sold at a higher used price.
3. The selfie camera has a positive correlation with used price and can be drive up the price of a used phone.
4. The phones with iOS operating system have a stable market price it has a positive correlation with used phone price.
5. The brands Blackberry and Google have positive correlation with Used price while brands OnePlus, Nokia and Infinix have negative correlation

Business Insights and Recommendations

1. ReCell should give promotions to phones that are highly used. The more this phone sits in the store, its value is going down day-by-day.
2. ReCell should try to increase the stock of 5G phones, and the market is pointing towards that direction. Also have some deals on phones without 4G or 5G capability so that they can be sold easily.
3. ReCell can build an inventory of phones with high resolution selfie cameras.
4. ReCell can invest in procuring more phones with iOS operating system and their price does not go down with age.
5. ReCell can build an inventory of phones from the brands Blackberry and Google.

greatlearning
Power Ahead

Happy Learning !