

Business Presentation ReCell Case

Business Problem Overview and Solution Approach

- Core business idea – ReCell (a startup company), aims to tap the potential of refurbished and used devices market. With the goal of using an ML-based solution to develop a dynamic pricing strategy for used and refurbished smartphones.
- Problem to tackle
 - the company wants to analyze the data provided,
 - build a linear regression model to predict the price of a used phone,
 - and identify factors that significantly influence it.

Problem to tackle (Cont'd) – Questions to answer

- What does the distribution of used phone prices look like?
- What percentage of the used phone market is dominated by Android devices?
- The amount of RAM is important for the smooth functioning of a phone. How does the amount of RAM vary with the brand?
- A large battery often increases a phone's weight, making it feel uncomfortable in the hands. How does the weight vary for phones offering large batteries (more than 4500 mAh)?
- Bigger screens are desirable for entertainment purposes as they offer a better viewing experience. How many phones are available across different brands with a screen size larger than 6 inches?
- Budget phones nowadays offer great selfie cameras, allowing us to capture our favorite moments with loved ones. What is the distribution of budget phones offering greater than 8MP selfie cameras across brands?
- Which attributes are highly correlated with the used phone price?

Data Overview

Variable	Description	Data Types
brand_name	Name of manufacturing brand	Object
os	OS on which the phone runs	Object
screen_size	Size of the screen in cm	Float64
4g	Whether 4G is available or not	Object
5g	Whether 5G is available or not	Object
main_camera_mp	Resolution of the rear camera in megapixels	Float64
selfie_camera_mp	Resolution of the front camera in megapixels	Float64
int_memory	Amount of internal memory (ROM) in GB	Float64
ram	Amount of RAM in GB	Float64
battery	Energy capacity of the phone battery in mAh	Float64
weight	Weight of the phone in grams	Float64
release_year	Year when the phone model was released	Int64
days_used	Number of days the used/refurbished phone has been used	Int64
new_price	Price of a new phone of the same model in euros	Float64
used_price	Price of the used/refurbished phone in euros	Float64

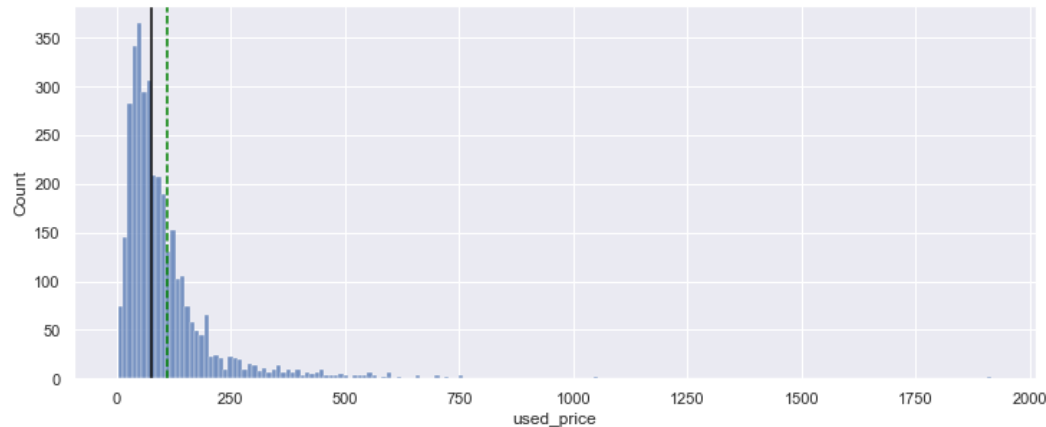
Observations	Variables
3,571	15

Notes:

- No duplicate records.
- Missing values were replaced with median values.
- Screen size metric convert to inches from cm.
- Outliers were not treated. We do not want to loose valuable information.
- Brand name, OS, 4G and 5G features were encoded.

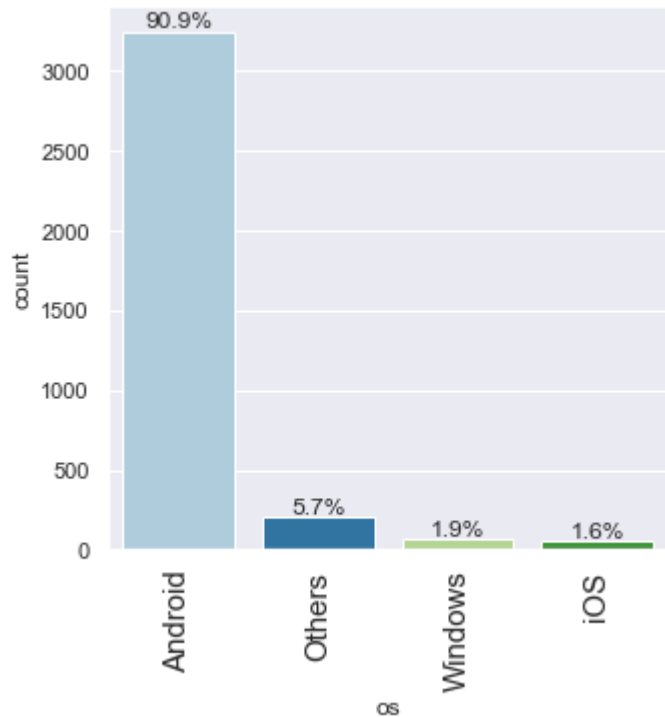
What does the distribution of used phone prices look like?

- Used Price is right skewed. There are a number of phones with high used prices.
- There are some outliers with more than 1,750 euros in used prices.
- Mean value of used price is about 110 euros.
- OnePlus, Apple and Google brands have the highest mean used prices with 318, 293, 239 euros respectively.
- Celkon, Spice and Karbonn have the lowest mean used prices with 22, 32, 38 euros respectively.



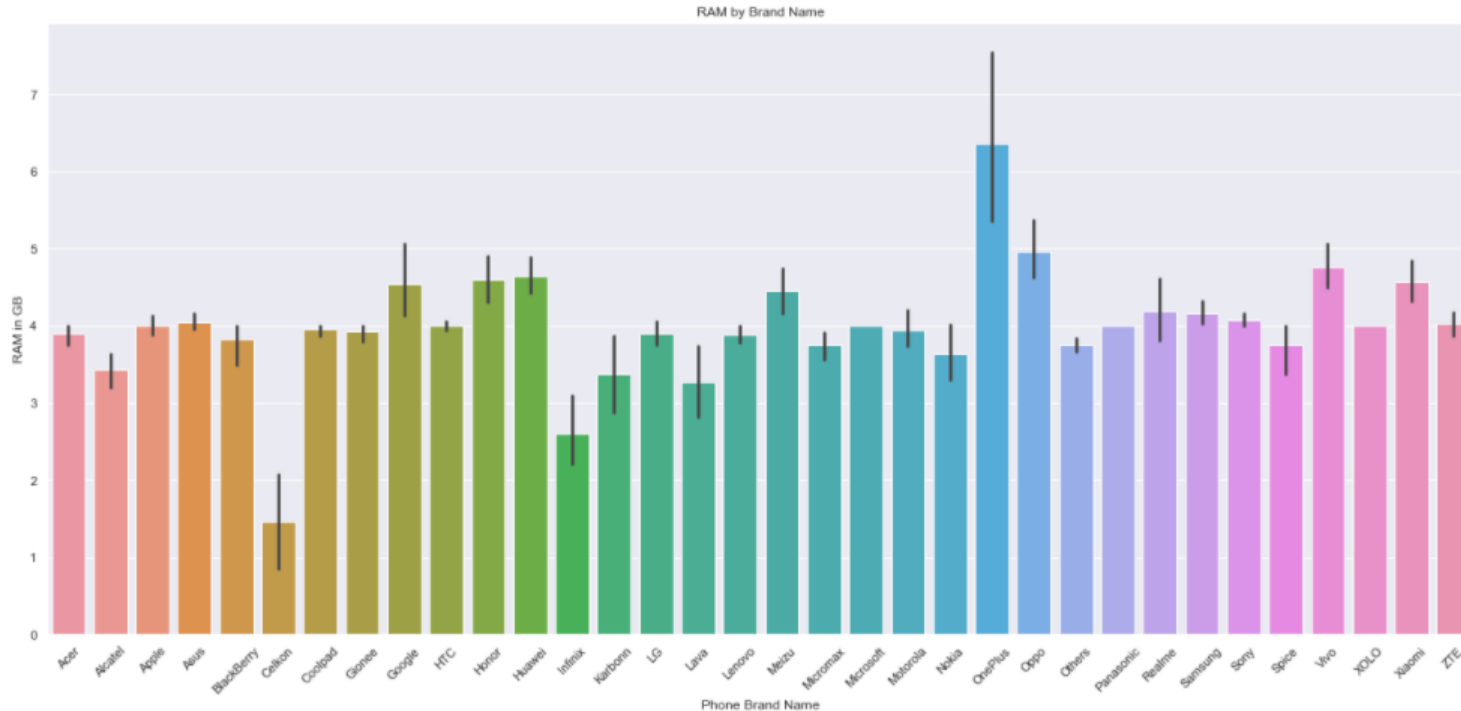
What percentage of the used phone market is dominated by Android devices?

- Android Operating system accounts for 90.9% of the market.



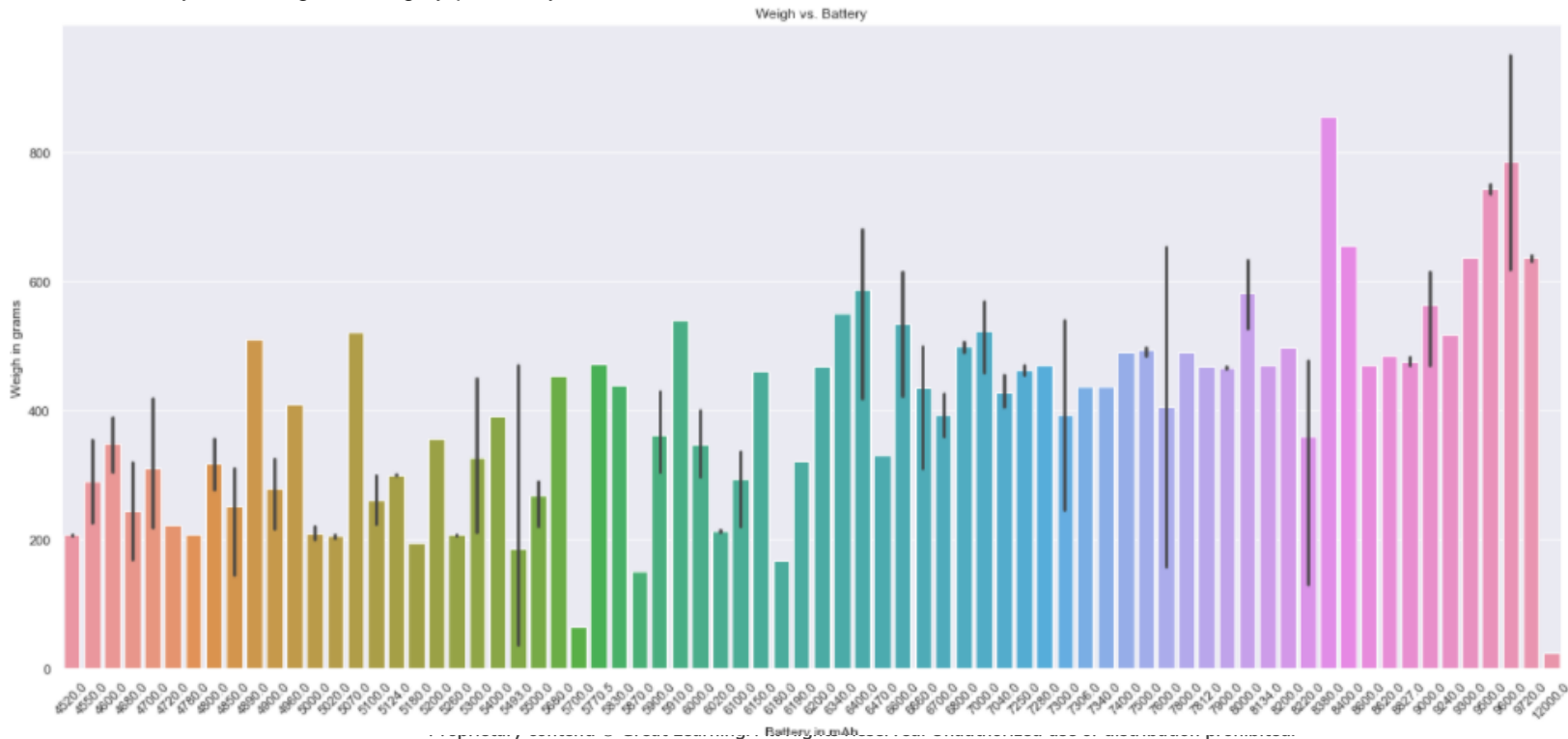
How does the amount of RAM vary with the brand?

- OnePlus has the highest average RAM (6.4GB) followed by Oppo (5GB) and Vivo (4.8GB).
- Celkon has the lowest average RAM (1.5GB) followed by Infinix (2.6GB) and Lava (3.3GB).
- Huawei and Nokia have the outliers with RAM of 16GB.
- Few brand names offer phones with only 4GB RAM with no outliers (Microsoft, Panasonic and XOLO).
- All brand names offer phones with 4GB R



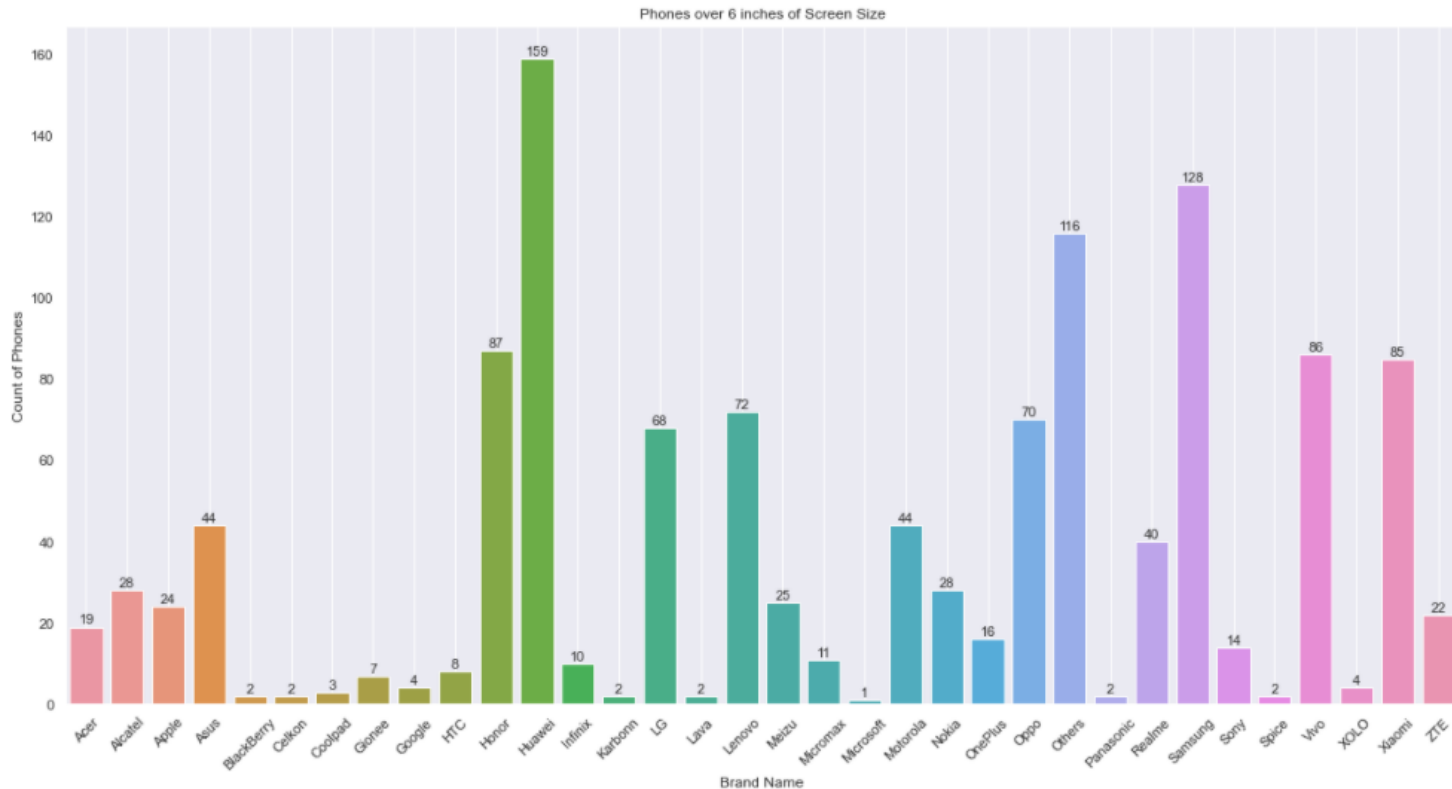
How does the weight vary for phones offering large batteries (more than 4500 mAh)?

- Google has the highest mean weight of all brands with 517 grams, followed Lenovo (460.5 grams) and Sony (439.5 grams).
- Most of the phones (206 units) have battery size within 4,500 to 5,500 mAh.
- Samsung has an outlier phone with the highest battery capacity (12,000 mAh) and lowest weight (3.17 grams).
- Battery and Weight are highly positively correlated



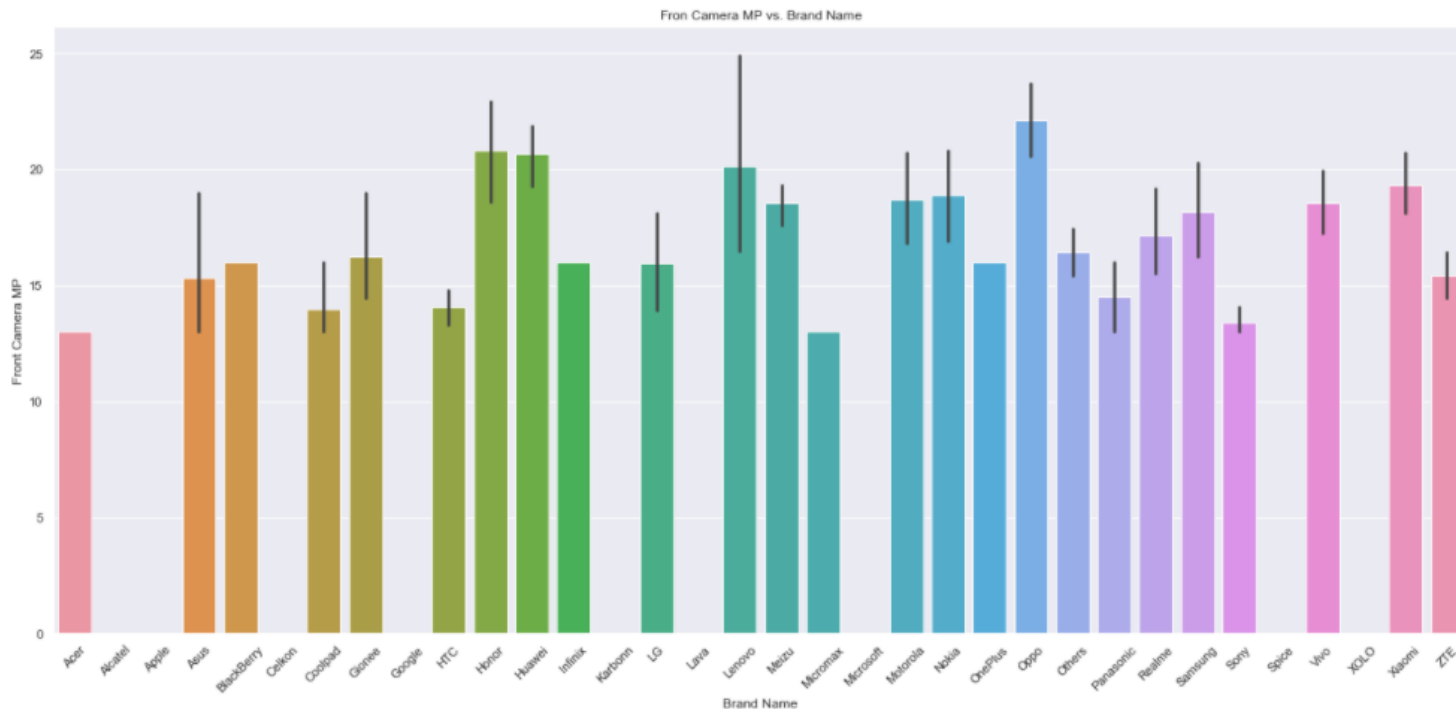
How many phones are available across different brands with a screen size larger than 6 inches?

- There are 1,235 phones with screen size greater than 6 inches and these account for approximately 35% of all used phones.
- Huawei has the most phones with over 6 inches of screen in this population. This is approximately 13% of dataset.



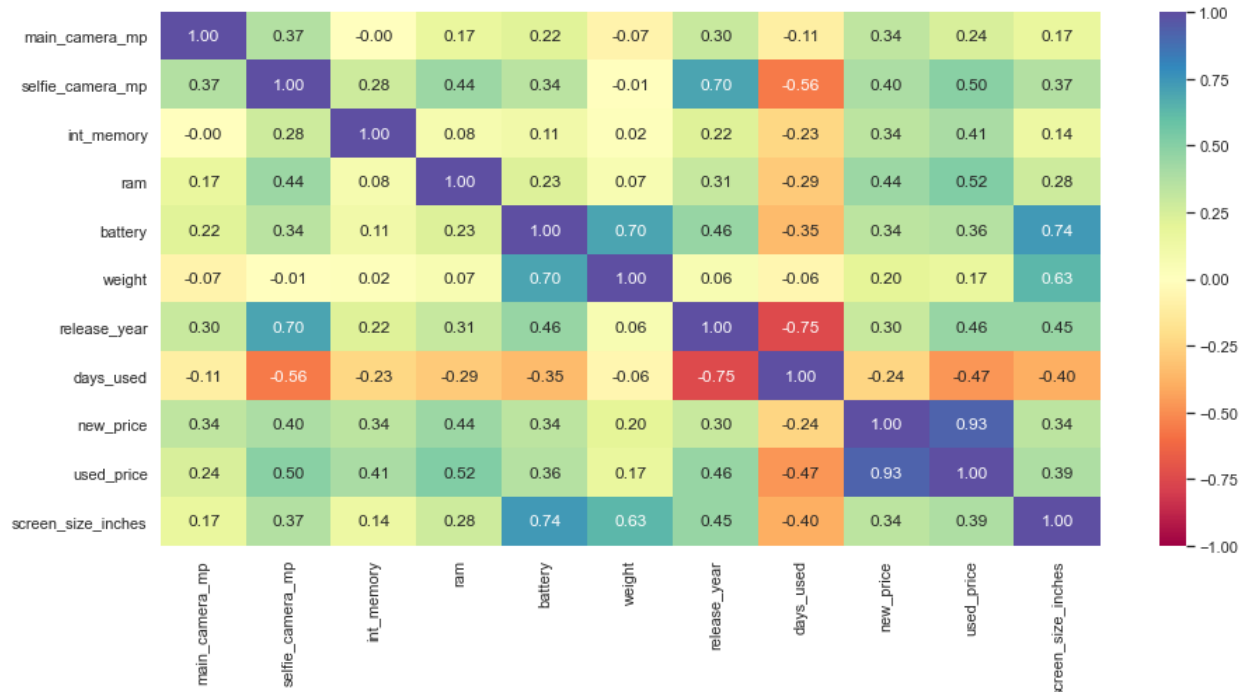
What is the distribution of budget phones offering greater than 8MP selfie cameras across brands?

- There are 666 phones with Selfie cameras with more than 8 megapixels and these account for approximately 19% of all used phones.
- Oppo, Honor and Huawei are the brands with the highest average megapixels respectively with 22.1, 20.8 and 20.7 megapixels respectively.



Which attributes are highly correlated with the used phone price?

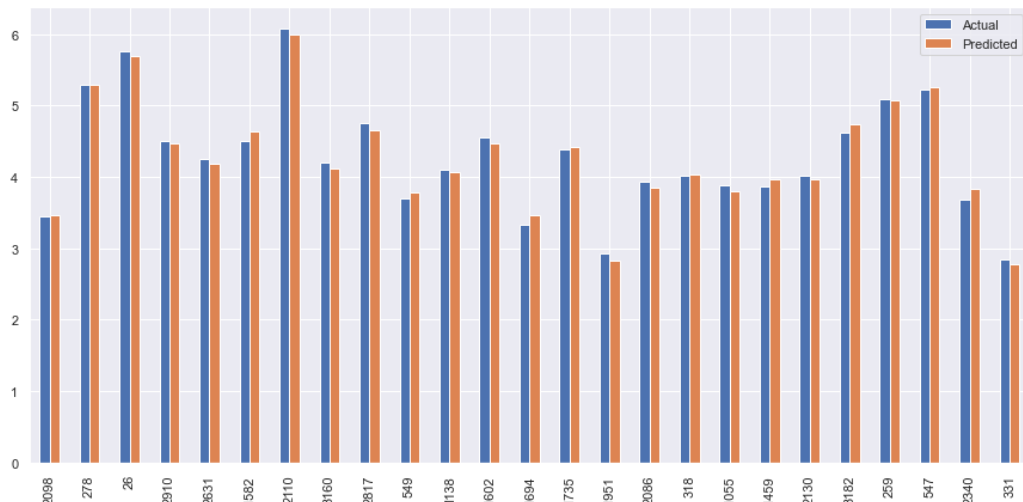
- *Used Price* is negatively correlated with *Days Used*. The more days used, the lower the used price.
- *Used Price* is highly positively correlated with *New Price*. Phones that costs more new and will most likely fetch higher resale price.
- *Used Price* is also positively correlated to all other variables (besides days used), with the top 3 being: *RAM*, *Selfie Camera mp* and *Release Year*. Higher RAM and Megapixels will fetch higher used resale price.
- *Battery* is positively correlated with *Screen Size* and *Weight*. Larger battery capacity is needed to support a larger screen size and larger battery weighs more.
- *Release Year* is positively correlated with the mp of the *selfie camera*. Newer camera has more megapixels for the selfie camera.



Model Performance Summary – Linear Regression

- To predict the used price (used_price_log), we split data into train and test dataset. Then built a linear regression model.
- We concluded that the model is good for prediction as well as inference purposes.
- The model is able to explain ~99% of the variation in the data, which is very good.
- The train and test RMSE and MAE are low and comparable. The model is not suffering from overfitting.
- The MAPE on the test set suggests we can predict within 1.7% of the used prices.

Performance	RMSE	MAE	R-squared	Adj. R-squared	MAPE
Train	0.082	0.068	0.990	0.990	1.668
Test	0.080	0.069	0.990	0.990	1.651



Model Performance Summary (Cont'd)

- Linear Regression Statsmodel performed slightly better than Linear Regression Sklearn
- Significant Features are - 'selfie_camera_mp', 'release_year', 'days_used', 'new_price_log', 'brand_name_Gionee', 'brand_name_Lenovo', '4g_yes', '5g_yes'

OLS Regression Results

Dep. Variable:	used_price_log	R-squared:	0.990			
Model:	OLS	Adj. R-squared:	0.990			
Method:	Least Squares	F-statistic:	3.088e+04			
Date:	Sun, 24 Oct 2021	Prob (F-statistic):	0.00			
Time:	05:57:34	Log-Likelihood:	2702.0			
No. Observations:	2499	AIC:	-5386.			
Df Residuals:	2490	BIC:	-5334.			
Df Model:	8					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]

const	-10.3593	2.856	-3.627	0.000	-15.959	-4.759
selfie_camera_mp	0.0009	0.000	2.399	0.017	0.000	0.002
release_year	0.0051	0.001	3.616	0.000	0.002	0.008
days_used	-0.0011	1.03e-05	-106.061	0.000	-0.001	-0.001
new_price_log	0.9847	0.003	347.521	0.000	0.979	0.990
brand_name_Gionee	-0.0381	0.013	-2.825	0.005	-0.065	-0.012
brand_name_Lenovo	-0.0162	0.008	-2.035	0.042	-0.032	-0.001
4g_yes	-0.0120	0.005	-2.497	0.013	-0.021	-0.003
5g_yes	0.0324	0.009	3.434	0.001	0.014	0.051
=====						
Omnibus:	193.242	Durbin-Watson:	2.036			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	320.503			
Skew:	-0.576	Prob(JB):	2.53e-70			
Kurtosis:	4.324	Cond. No.	3.69e+06			

Linear Regression sklearn Linear Regression statsmodels

RMSE	0.081557	0.082072
MAE	0.067524	0.068225
R-squared	0.990146	0.990021
Adj. R-squared	0.989953	0.989985
MAPE	1.650207	1.667946

Business Insights and Recommendations

Conclusion

- *days_used*, *new_price*, *brand_name_Gionee*, *brand_name_Lenovo*, *selfie_camera_mp*, *4g* and *5g* are the factors that most affects the price of used phones.
- A higher priced new phone will lead to an increase in the price of used phone.
- *days_used* and *brand_name_Gionee*, *brand_name_Lenovo*, and *4g* have a negative effect on the price of used phones.
- *new_price*, *selfie_camera_mp*, *release_year* and *5g* have positive coefficients. So, as they increase, used price also increases.

Recommendation

- Focus on newer phones with high selfie camera resolution and 5g. There have a favorable impact on used phone prices.
- Avoid older phones with 4g and