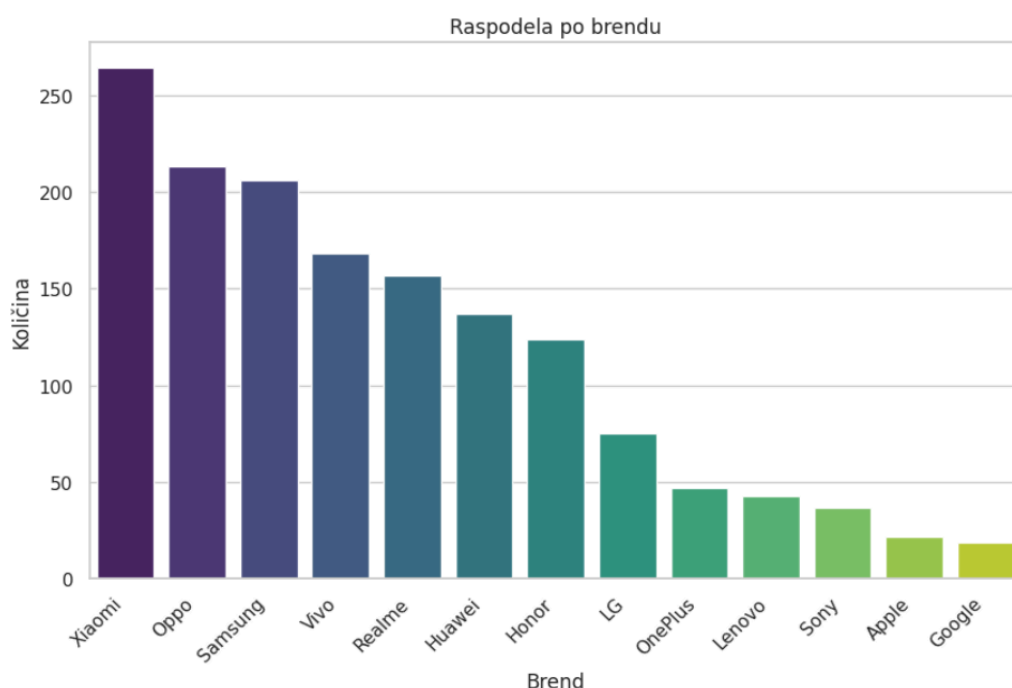


# Eksplorativna analiza podataka

## Brend

### Distribucija telefona prema brendu

Kao početna tačka predstavke podataka nad kojima se vrši analiza, obučavanje i provera algoritama za procenu cene telefona po njegovim karakteristikama, uzeta je podela mobilnih telefona po brendovima kojim pripadaju. U obrađivanom skupu podataka, postoji 13 brendova i to: Xiaomi, Oppo, Samsung, Vivo, Realme, Huawei, Honor, LG, OnePlus, Lenovo, Sony, Apple i Google, respektivno po broju modela.



*Grafikon 1: Raspodela brendova telefona*

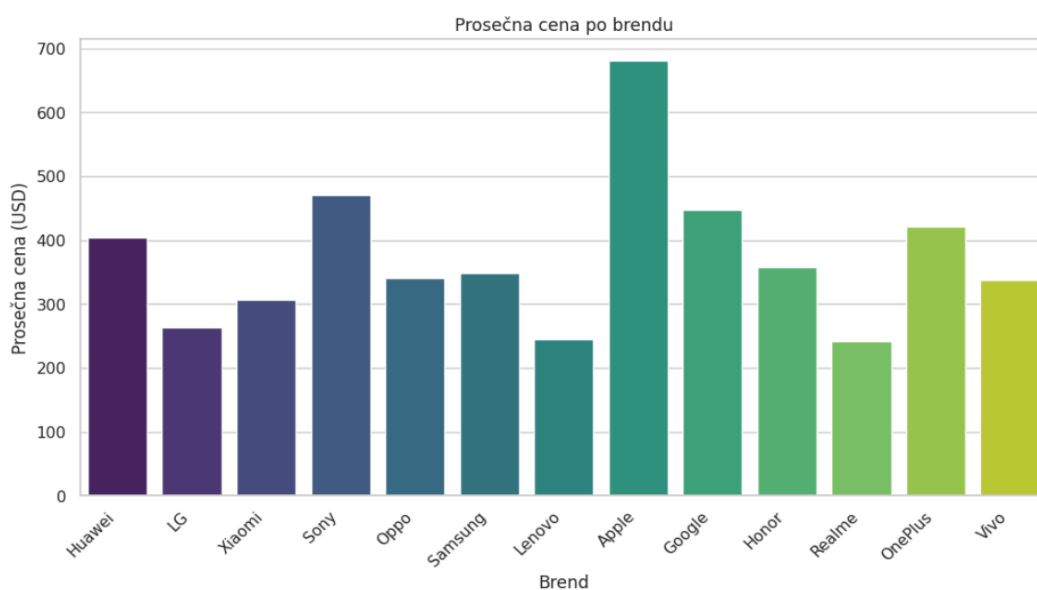
Na grafikonu 1 je grafički prikazana raspodela telefona prema brendu, gde prva tri mesta zauzimaju Xiaomi, Oppo i Samsung, dok Apple zauzima 12. mesto, što je važno istaći jer taj brend odstupa od ostalih zbog cenovne kategorije, jer ne postoje niskobudžetni telefoni.

➡ količina telefona po brendu:

	Brend	Količina	brand
0	NaN	264	Xiaomi
1	NaN	213	Oppo
2	NaN	206	Samsung
3	NaN	168	Vivo
4	NaN	157	Realme
5	NaN	137	Huawei
6	NaN	124	Honor
7	NaN	75	LG
8	NaN	47	OnePlus
9	NaN	43	Lenovo
10	NaN	37	Sony
11	NaN	22	Apple
12	NaN	19	Google

Ukupna količina svih telefona je: 1512

## Odnos cene i brenda



*Grafikon 2: Raspodela prema prosečnoj ceni određenog brenda*

Kao što je ukazano ranije a primetno na grafikonu 2, Apple telefoni u proseku koštaju najviše prema ostatku brendova, iako ih ima skoro najmanje brojevno. Zbog svoje visoke cene i za najjeftinije modele, prave veliko odstupanje od ostalih brendova koje varira od 41 do 183 posto. Cene su prikazane u američkim dolarima (USD).

## Operativni sistem

### Distribucija telefona prema operativnom sistemu

U daljem postupku analize skupa podataka, izvršeno je filtriranje i pojednostavljivanje podataka prema operativnom sistemu koji se koristi, na način da su podverzije i posebni

Android 11: 322,  
Android 10: 318,  
Android 12: 236,  
Android 9: 212,  
Android 13: 154,  
Android 8: 145,  
Android 7: 78,  
Android 6: 11,  
EMUI 12: 6,  
iOS 15: 5,  
iOS 13: 4,  
iOS 16: 4,  
iOS 14: 4,  
Android: 3,  
iOS 12: 3,  
Android 5.1: 2,  
iOS 11: 2,  
EMUI 13: 2,  
Tizen 3.0: 1 uređaj.

### Cena telefona prema sistemu

Box plot showing the distribution of phone prices (Cena (USD)) across different operating systems (Operativni Sistem).

The Y-axis represents the price in USD, ranging from 0 to 2000. The X-axis lists the operating systems: Android, Tizen, iOS, and EMUI.

The plot displays the median (horizontal line inside the box), the interquartile range (the box itself), and the range of the data (whiskers). Outliers are shown as individual points.

Approximate values for the box plots:

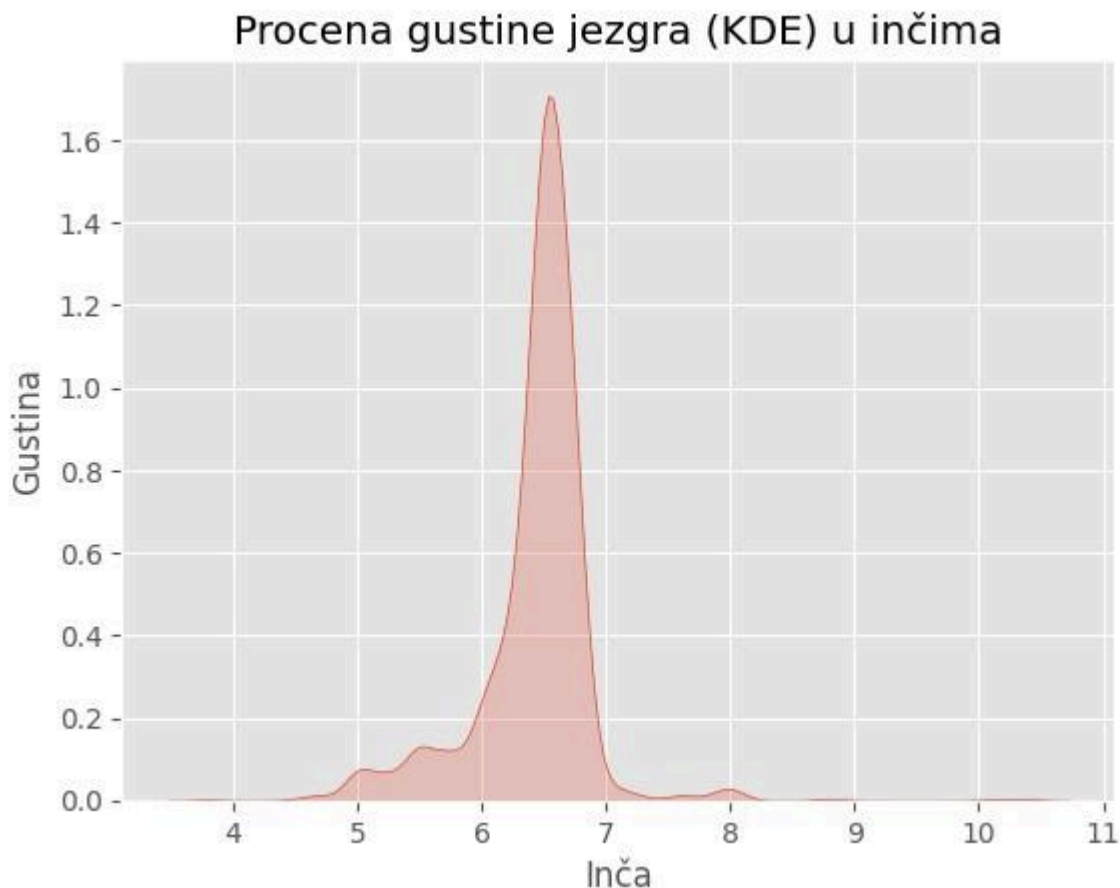
Operativni Sistem	Min (Whisker)	Q1 (Box Bottom)	Median (Line)	Q3 (Box Top)	Max (Whisker)	Outliers (USD)
Android	~50	~200	~250	~400	~750	~1700, ~1800, ~1900, ~2000, ~2100, ~2200, ~2300
Tizen	~250	~250	~250	~250	~250	None
iOS	~200	~400	~600	~900	~1600	None
EMUI	~200	~250	~250	~300	~300	None

*Grafikon 3: Raspodela prema odnosu cene i operativnog sistema*

Ukoliko se posmatra odnos cene telefona prema operativnom sistemu koji koristi, a to je grafički prikazano a grafikonu 3, može se izvesti zaključak da najskuplji telefoni pokreću Android operativni sistem, ali u konačnici, iOS u proseku ima najskuplje modele, čemu doprinosi njihova visoka cena i mali broj različitih uređaja.

## Ekran

### Veličina ekrana

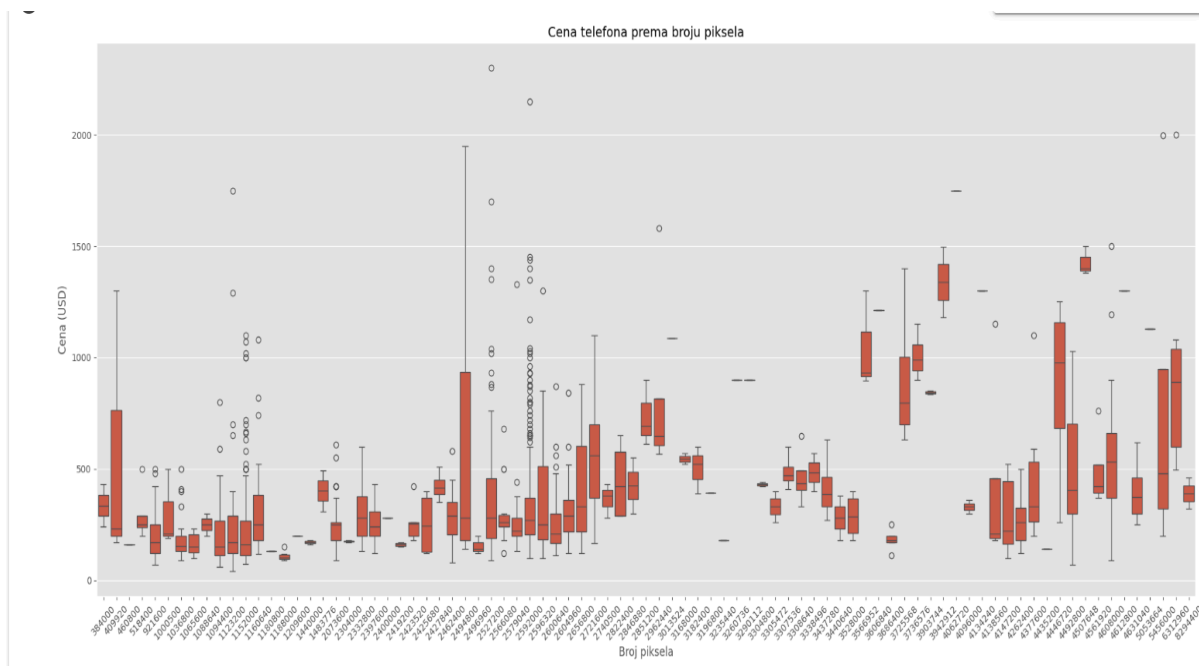


*Grafikon 4: Raspodela telefona prema veličini ekrana u inčima*

Ukoliko se primeni statistička tehnika KDE (Kernel Density Estimate) na skup podataka, gde se kao ulazni parametar koristi veličina ekrana u inčima, dobija se prikaz kao u grafikonu 4. Iz navedenog grafikona, zaključuje se da je najveći broj telefona u rasponu od 6 do 7 inča dijagonale ekrana.

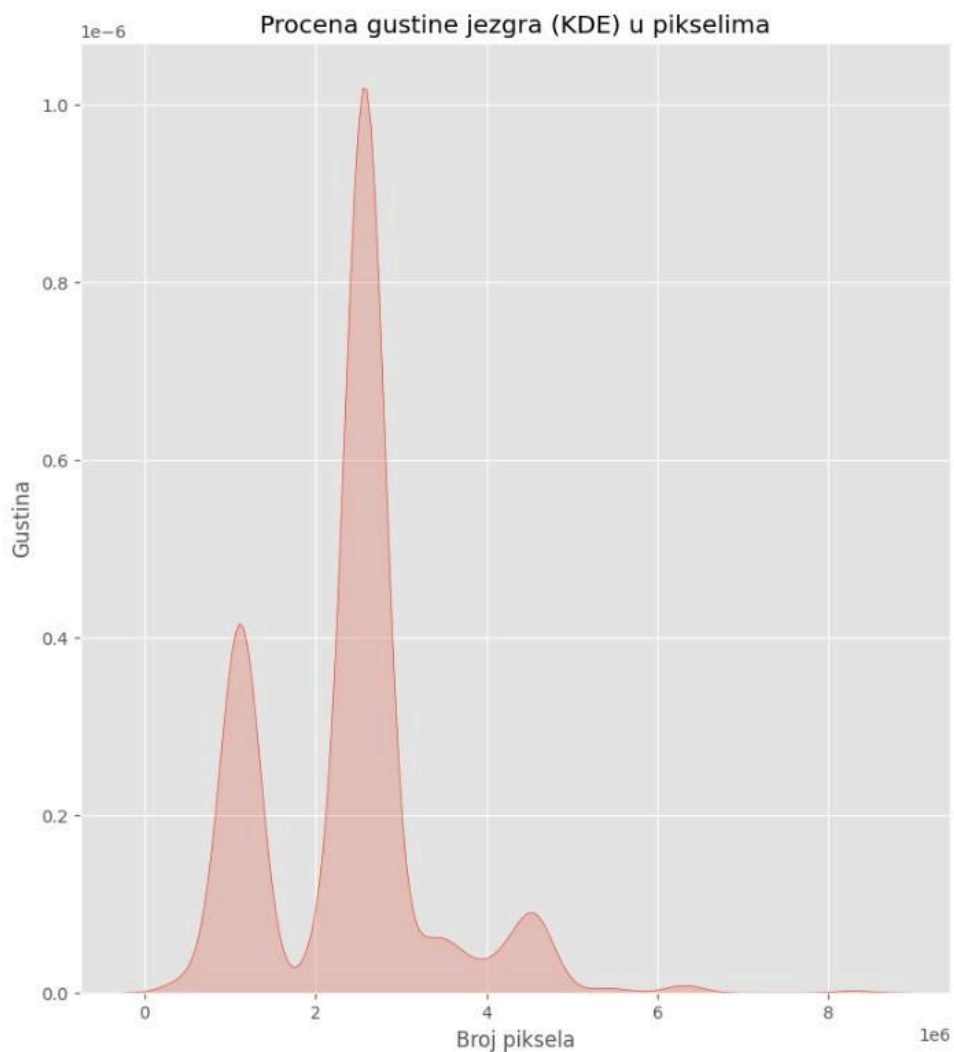
### Rezolucija ekrana

Kako bi bilo moguće izračunati i predstaviti telefone prema rezoluciji ekrana koju poseduju, neophodno je primeniti posebnu funkciju za prilagođavanje podatka o rezoluciji, na način da se na primer rezolucija izražena u pikselima 720x1080 i zaista pomnoži, što daje 777,600 piksela. Ista funkcija se primenjuje na svaku jedinicu skupa podataka, kreira se nova kolona sa ukupnim brojem piksela, a kolona za rezoluciju u starom formatu se uklanja.



**Grafikon 5: Raspodela telefona prema broju piksela i ceni**

Na grafikonu 5 je prikazana raspodela mobilnih telefona u odnosu na broj piksela i njegovu cenu.



*Grafikon 6: Raspodela telefona prema količini piksela*

Sa grafikona 6 primenom statističke tehnike KDE dobija se podatak da najveći broj telefona sadrži broj piksela koji varira od  $2 \cdot 10^6$  do  $3 \cdot 10^6$ . Primetan je i značajan broj uređaja sa brojem piksela od oko  $1 \cdot 10^6$ .

## Kamera

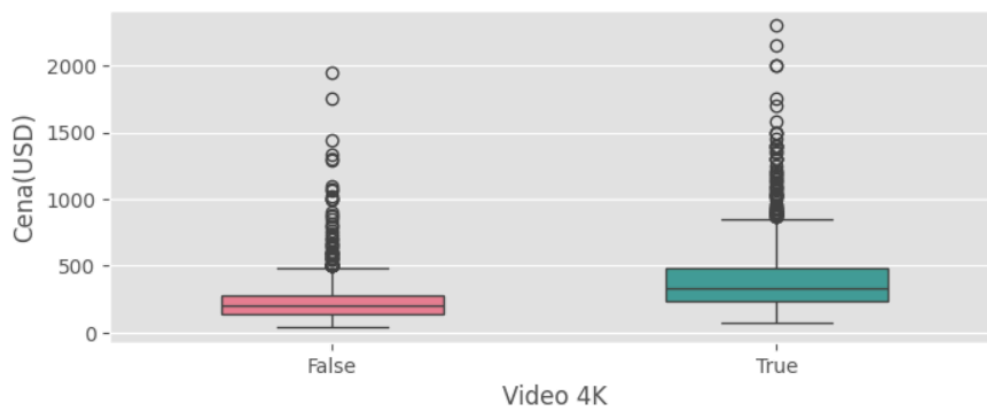
### Rezolucija video snimka

Kako postoji veliki broj atributa koji se odnose na kamere, izdvojeni su samo oni atributi za koje je zaključeno da najviše utiču na cenu mobilnog telefona.

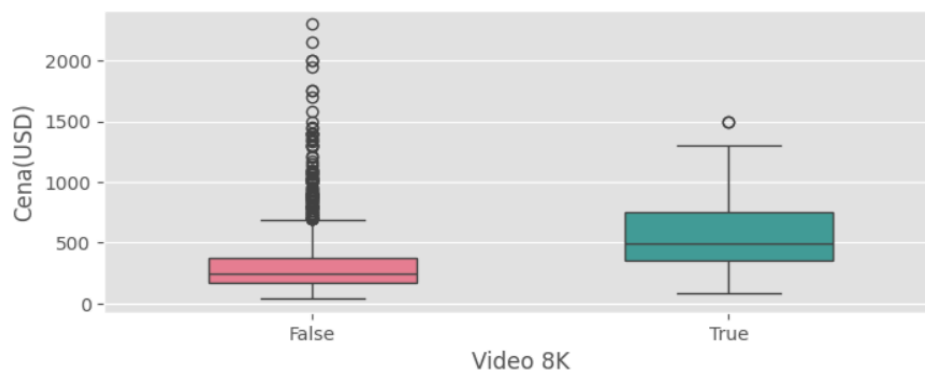


*Grafikon 11: Raspodela telefona prema tome da li podržavaju 720p snimanje*

5: `plt.boxplot([Cena(USD) for Video 4K in [False, True]])`



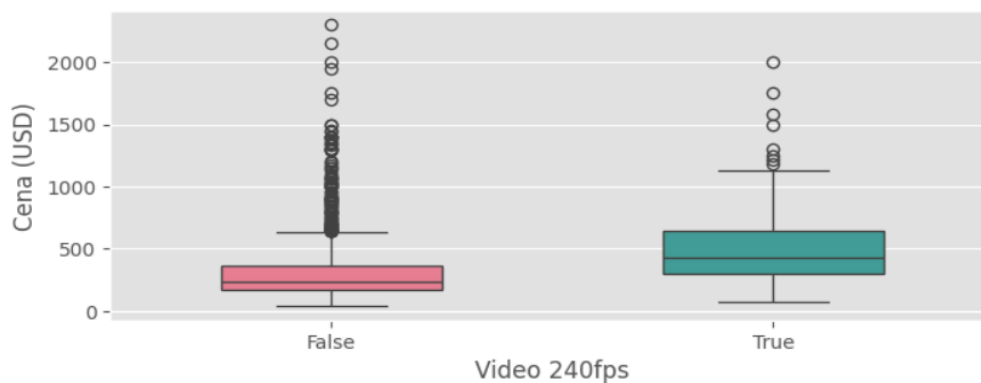
*Grafikon 12: Raspodela telefona prema tome da li podržavaju 4K snimanje*



*Grafikon 13: Raspodela telefona prema tome da li podržavaju 8K snimanje*

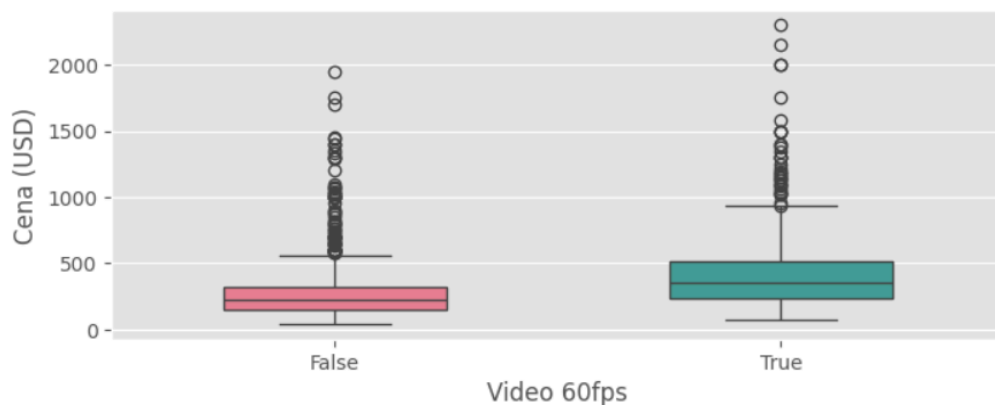
Najveći broj telefona iz skupa podataka ne podržava snimanje u 8K rezoluciji, ali oni koji podržavaju imaju više cene (grafikon 13). Isto se odnosi i na telefone koji imaju podržano snimanje u 4K rezoluciji i njihova cena je viša u odnosu na one koji ne podržavaju (grafikon 12). Kao i kod prethodna dva, isti podatak se dobija i za 720p video snimke. Zaključak koji se može izvesti jeste da ova tri atributa direktno utiču na cenu telefona.

### Okvira po sekundi (FPS)



*Grafikon 14: Raspodela telefona prema tome da li podržavaju snimanje u 240fps*

Kako svega 203 mobilna telefona podržavaju pravljenje video snimaka u 240fps uz podatak da takvi telefoni imaju drastično veće cene od ostalih, to trening modelu ukazuje da izuzetno utiče na cenu mobilnog telefona.

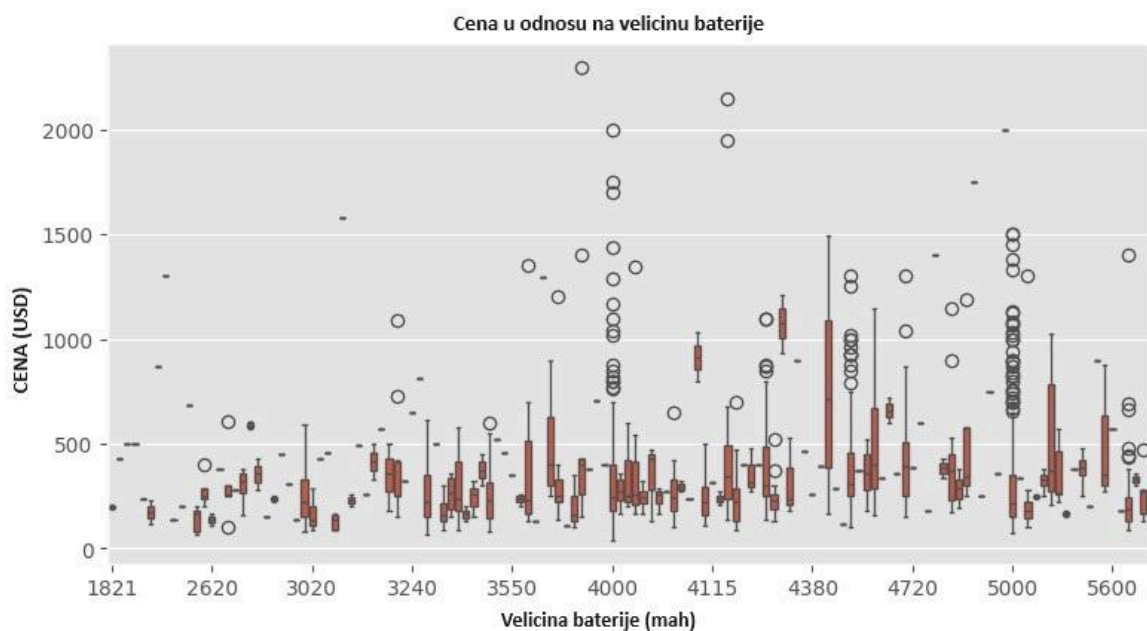


*Grafikon 15: Raspodela telefona prema tome da li podržavaju snimanje u 60fps*

Slično kao i za snimanje u 240fps, zaključeno je da snimanje u 60fps povećava krajnju cenu telefona. Napravljeno je sabiranje redova, odnosno sabirali su se 'TRUE' i 'FALSE' kako bi dobili novu kolonu 'better\_cam' i sve kolone koje imaju 'video' u nazivu su dropovane.

## Baterija

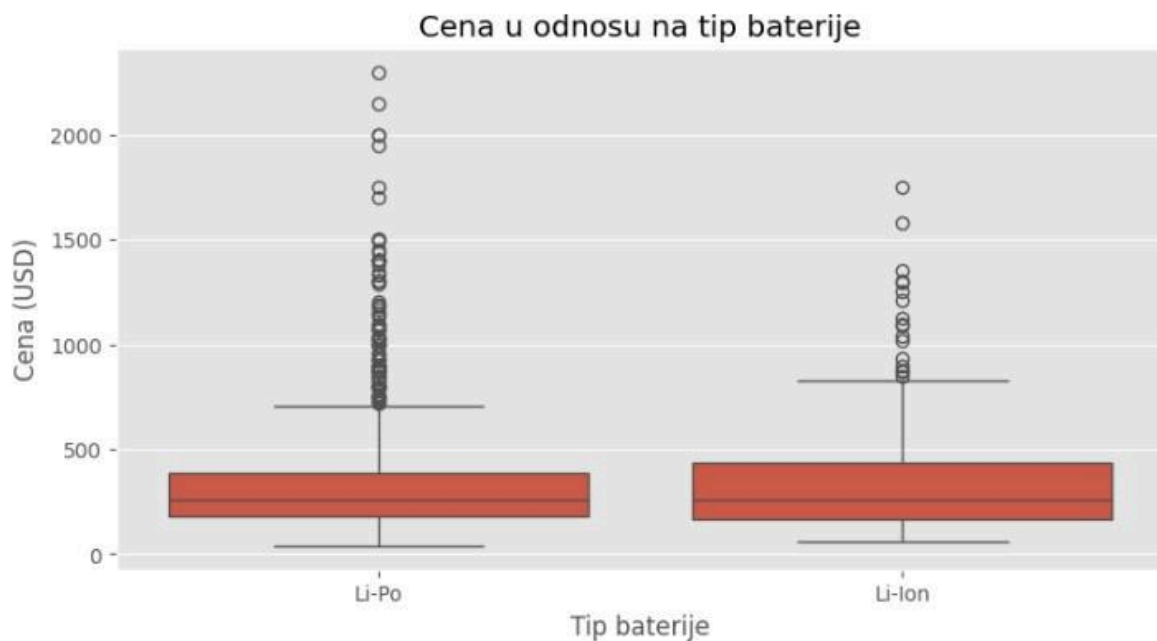
### Veličina baterije



*Grafikon 7: Raspodela telefona prema odnosu baterije i cene*



## Tip baterije

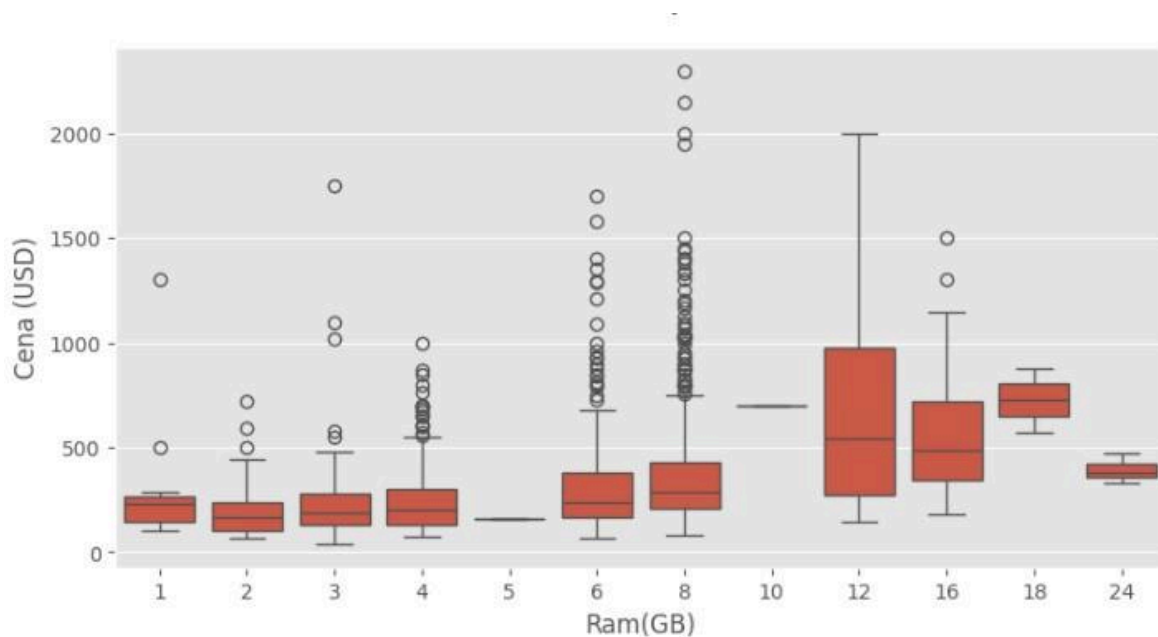


*Grafikon 8: Raspodela telefona prema odnosu tipa baterije i cene*

Prilikom obrade skupa podataka, primećeno je da telefoni poseduju dva tipa baterija a to su Li-Po i Li-Ion baterije. Analizirajući podatke, dolazi se do zaključka da skuplji telefoni u sebi sadrže Li-Po baterije, ali da veći broj telefona u proseku ima Li-Ion bateriju, što se može videti na grafikonu 8.

## Memorija

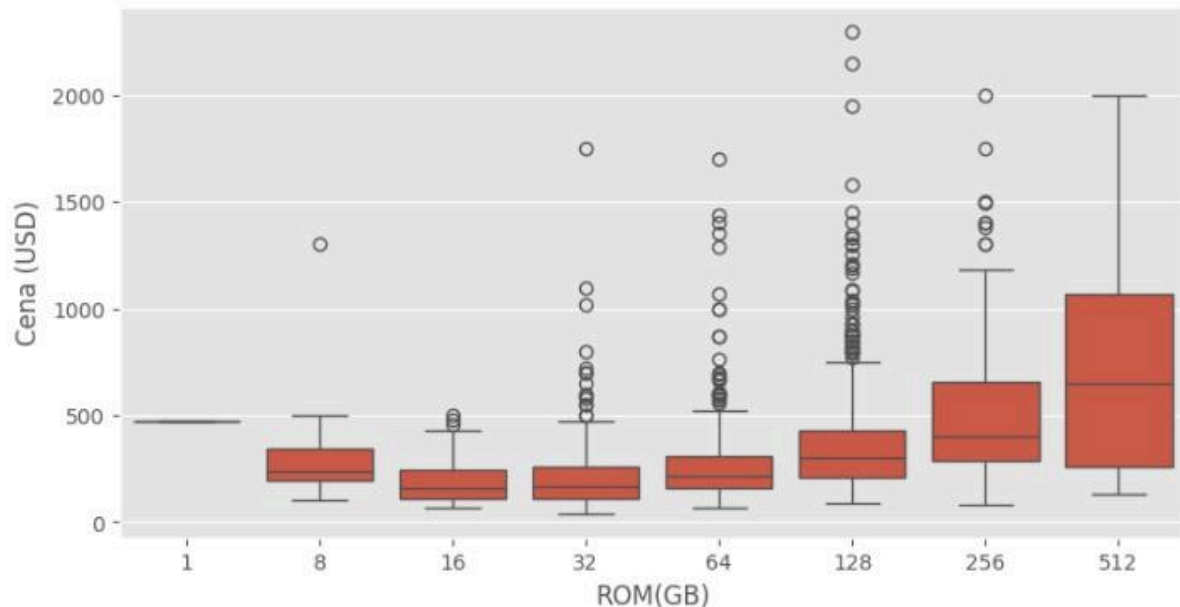
### RAM memorija



*Grafikon 9: Raspodela telefona prema odnosu količine RAM i cene*

Analizirajući količinu RAM memorije koju poseduju telefoni dolazi se do nekoliko zaključaka. Skuplji telefoni uglavnom sadrže 12 ili 16GB memorije, dok su 6 i 8GB prosečne količine memorije koje najveći broj telefona ima. Samim tim, količina RAM memorije direktno utiče na cenu mobilnog telefona.

## ROM memorija



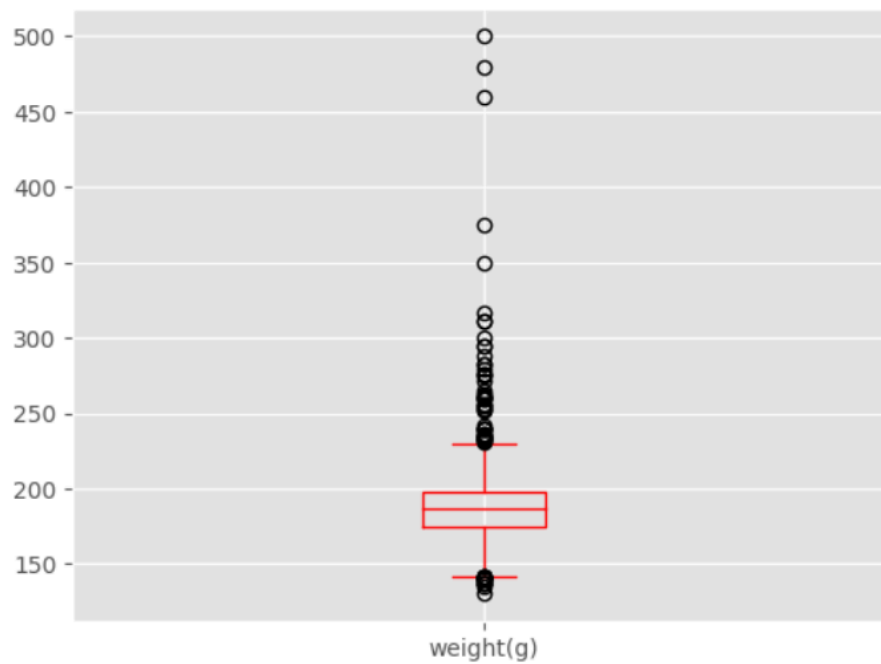
*Grafikon 10: Raspodela telefona prema odnosu količine skladišta i cene*

U nastavku analize podataka, izvršen je prikaz u vidu grafikona 10, gde su navedene uobičajene količine unutrašnje memorije koje telefoni poseduju. Najzastupljenija količina skladišne memorije je 128GB (677 jedinica) a onda sledi 64GB (382 jedinica). Sa porastom količine unutrašnje memorije raste i prosečna cena uređaja što dovodi do zaključka da ona direktno utiče na cenu mobilnog telefona.

## Outlejeri

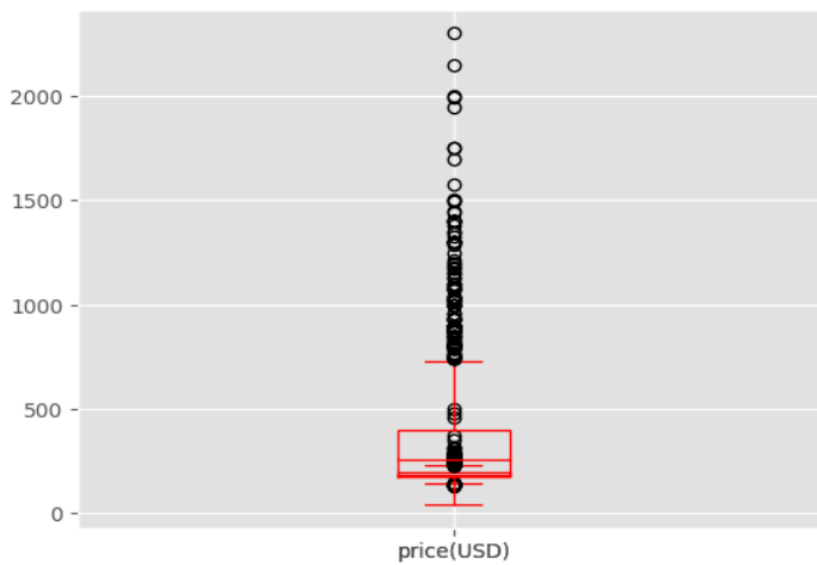
Sto se tice vrednosti koje izlaze iz opsega,  $\text{weight(g)} > 450$

<Axes: >

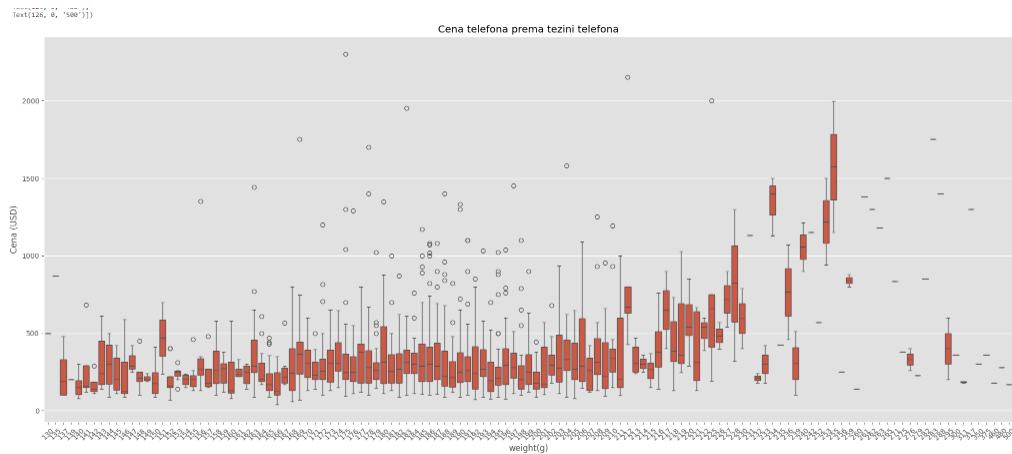


$\text{price(USD)} > 200$

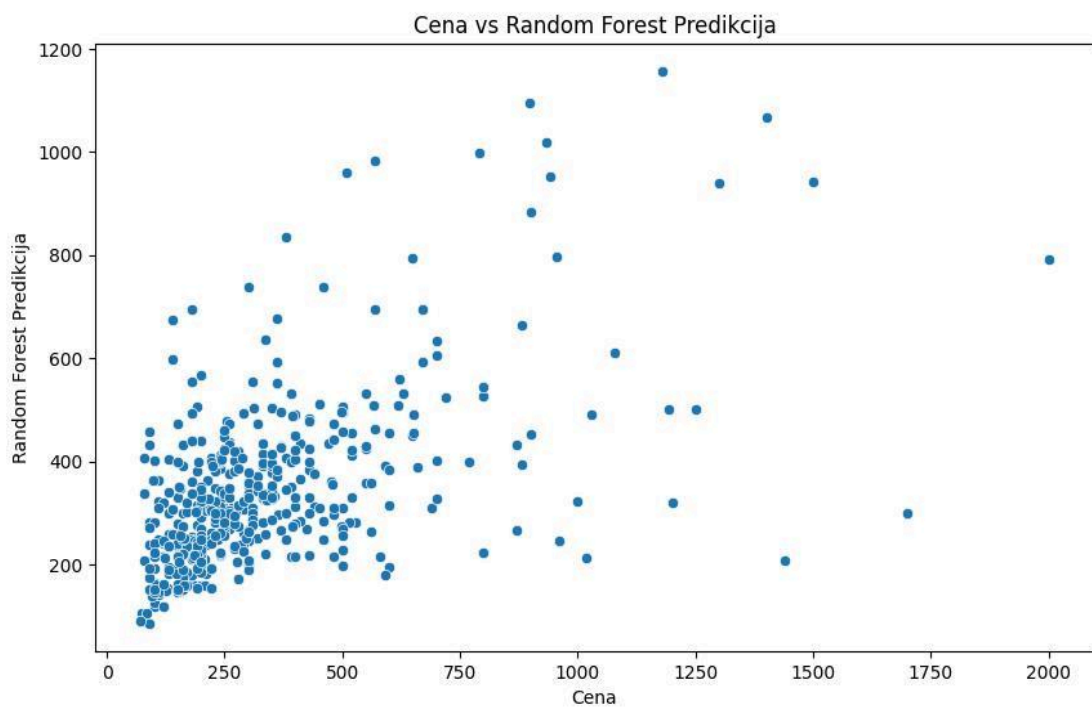
<Axes: >



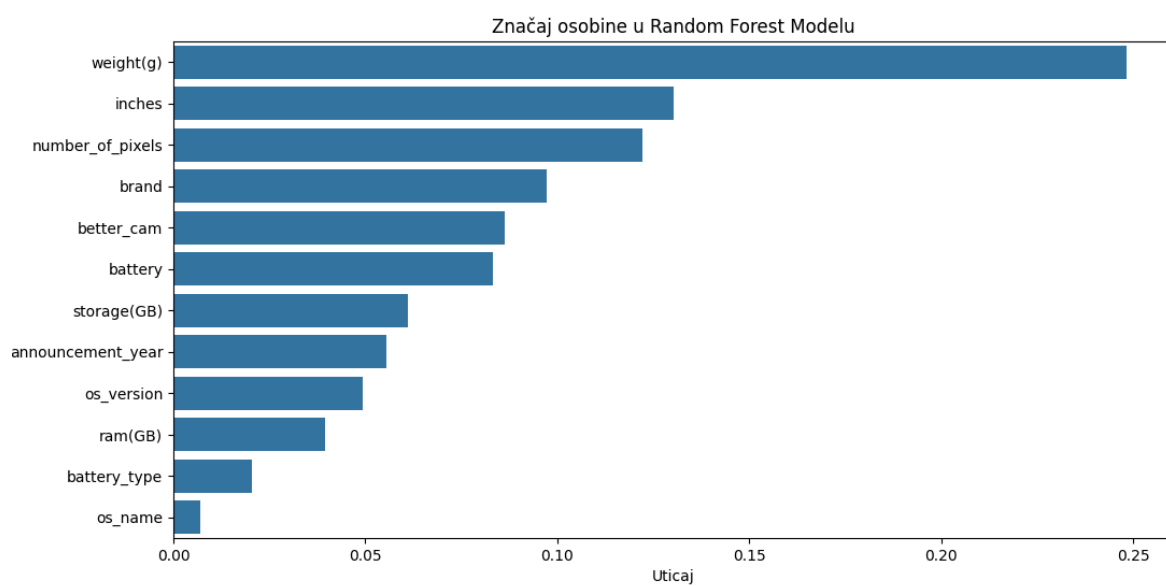
Cena prema tezini telefona:



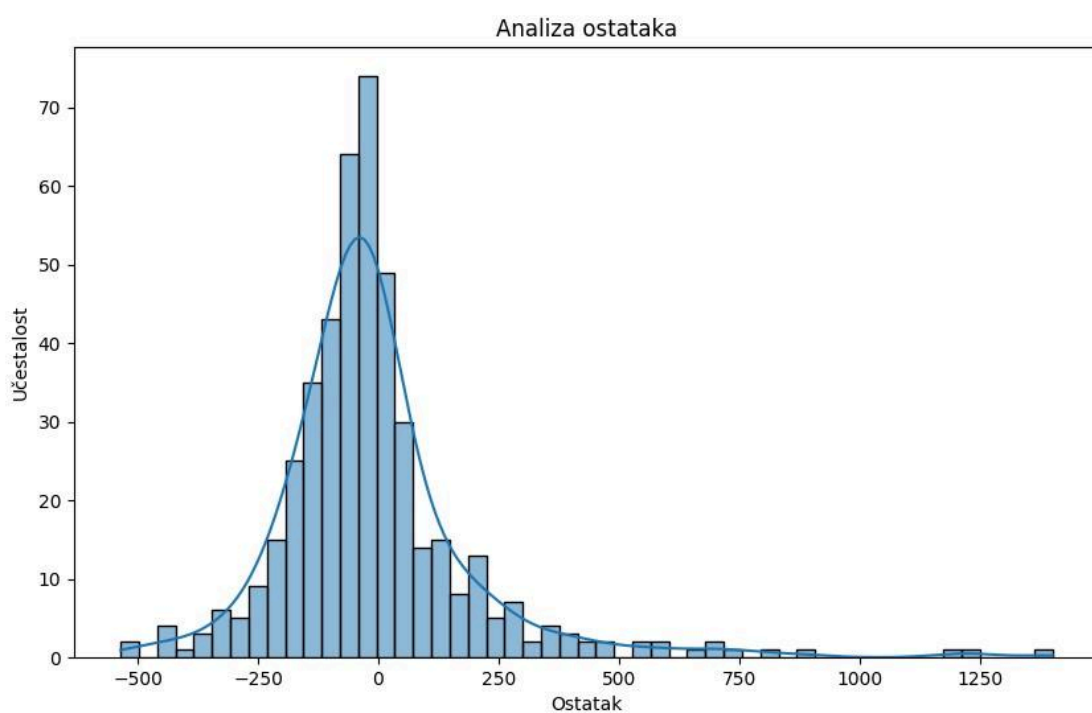
Grafički prikaz aktuelne cene i cene nastale primenom Random Forest predikcije koji podrazumeva da tačke bliže dijagonali ( $x=y$ ) predstavljaju tačnu predikciju.



Grafički prikaz značaja osobina mobilnog telefona odnosno njihovog uticaja na cenu, a na osnovu tih informacija su vršena predviđanja.



Predstavka ostataka pri oduzimanju stvarne cene i predviđene cene.



Grafički prikaz poređenja ostataka i stvarnih cena telefona je ispod

