Statystyczna Analiza Danych

Analiza skupień i porządkowanie liniowe wybranych modeli telefonów

# Wstęp

Projekt ma za zadanie przedstawienie analizy różnych modeli telefonów dostępnych na rynku. Dane techniczne zostały pozyskane ze strony mgsm.pl, a ich pojedyncze braki uzupełnialiśmy z ofert sprzedaży znajdujących się na allegro.pl. Ceny zostały zaczerpnięte ze strony ceneo.pl. Całość została zrealizowana w języku python. Projekt składa się z dwóch części: porządkowania liniowego oraz analizy skupień.

# Przedstawienie danych

## Ogólnie o modelach i metodyce ich doboru

Zebraliśmy 23 modele telefonów różnych producentów, takich jak: Apple, Samsung, Microsoft, Motorola, Google, Huawei, Asus, Nokia, HTC, Xiaomi, Oppo. Nie posiadaliśmy żadnej reguły doboru modeli telefonów. Wykorzystaliśmy te dane, które nas ciekawiły, lub dodaliśmy pierwszy lepszy popularnego producenta, którego uznaliśmy, że powinien otrzymać jakiegoś reprezentanta. Między innymi byliśmy zainteresowani flagowymi produktami firmy Apple (czyli na chwilę obecną iPhonem 15 w wersji zwykłej jak i pro max), a także starszym modelem owej firmy (iPhone 13). Wykorzystaliśmy wiele niszowych modeli, takich jak ULEFONE NOTE 16 PRO. Nie umknęły nam także starsze modele np. HTC 10. Udało nam się także znaleźć zmodernizowaną wersję telefonu Nokii 3310.

Po prawej stronie można zobaczyć dokładne modele, które zostały wykorzystane.

## Wybrane parametry

Ze względu na zróżnicowane zapotrzebowania konsumentów wybraliśmy 9 parametrów, które według nas, są najważniejsze przy zakupie telefonu. Mianowicie:

* opinie użytkowników (w skali od 0 do 10)
* cena (w złotówkach polskich)
* wbudowana pamięć (w GB)
* pamięć RAM (w GB)
* waga (w gramach)
* pojemność baterii (w mAh)
* zegar procesora (w Mhz)
* data premiery (liczone jako różnica obecnego kwartału a kwartału wypuszczenia danego modelu)
* ilość megapikseli głównego aparatu

Zdajemy sobie sprawę z faktu, iż nie są to wszystkie parametry na które konsument zwraca uwagę, jednakże są to według nas wystarczająco godni reprezentanci danych, różnych typów konsumentów. Jednocześnie już tak duża ilość parametrów sprawiła nam pewne kłopoty, gdyż nie każdy telefon ma wymienione wszystkie parametry. Najcięższym przypadkiem który zachowaliśmy to Nokia 3310 z 2017r. Owa modernizacja legendarnego już telefonu posiada oryginalny procesor którego nie można w taki sposób mierzyć. Z tego też powodu zastosowaliśmy tam wartość 0. Pozostałe modele telefonów, które nie posiadały wszystkich danych zostały usunięte i nie są nawet prezentowane w załączonym pliku z danymi.

## Dane

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Telefon\zmienne | Ocena | Cena | Pamięć wbudowana | Pamięć RAM | Waga (g) | bateria (mAh) | Zegar procesora (Mhz) | różnica kwartałów | Megapiksele |
| Apple iPhone 15 | 7,28 | 4192 | 128 | 6 | 171 | 3349 | 3460 | 1 | 48 |
| Samsung Galaxy A54 | 7,95 | 1599 | 128 | 8 | 202 | 5000 | 2400 | 3 | 50 |
| Motorola g72 | 6,89 | 899 | 128 | 8 | 166 | 5000 | 2200 | 4 | 108 |
| DOOGEE SMINI | 6,25 | 1200 | 256 | 8 | 155 | 3000 | 2200 | 0 | 50 |
| ULEFONE NOTE 16 PRO | 5 | 697,7 | 128 | 8 | 184 | 4400 | 1600 | 2 | 50 |
| Google Pixel 6pro | 7,51 | 2250 | 128 | 12 | 210 | 5003 | 2800 | 8 | 50 |
| HTC U23 PRO | 7,78 | 1899 | 256 | 12 | 205 | 4600 | 2400 | 2 | 108 |
| HTC DESIRE 22 PRO | 5,82 | 998 | 128 | 8 | 205,5 | 4520 | 2200 | 6 | 64 |
| HTC 10 | 8,78 | 555 | 32 | 4 | 161 | 3000 | 2200 | 30 | 12 |
| Xiaomi Redmi Note 13 Pro 5G | 6,81 | 3700 | 512 | 16 | 187 | 5100 | 2800 | 1 | 200 |
| SAMSUNG GALAXY S23 | 8,22 | 3048 | 128 | 8 | 167 | 3900 | 3360 | 3 | 50 |
| Huawei Mate 60 RS Ultimate | 6,79 | 11000 | 1000 | 16 | 246 | 5000 | 2620 | 1 | 48 |
| Apple iPhone 15 pro max | 8,22 | 9504 | 1000 | 8 | 221 | 4422 | 3770 | 1 | 48 |
| Apple iPhone 13 | 5,48 | 2819 | 128 | 4 | 174 | 4085 | 3220 | 9 | 12 |
| ASUS ROG PHONE 7 ULTIMATE | 6,83 | 6499 | 512 | 16 | 246 | 6000 | 4200 | 2 | 50 |
| ASUS ROG PHONE 7 | 8,23 | 3999 | 512 | 16 | 239 | 6000 | 3200 | 2 | 50 |
| POCO F3 | 7,47 | 1399 | 128 | 6 | 196 | 4520 | 3200 | 11 | 48 |
| OPPO A91 | 7,28 | 939 | 128 | 8 | 172 | 4025 | 2100 | 12 | 48 |
| POCO F5 PRO | 8,64 | 2369 | 256 | 12 | 204 | 5160 | 3200 | 2 | 64 |
| nubia red magic 7 pro | 6,07 | 4799 | 512 | 16 | 235 | 5000 | 3000 | 7 | 64 |
| NOKIA 3310 2017 DUAL SIM | 6,15 | 249 | 0,015625 | 0,015625 | 79,6 | 1200 | 0 | 26 | 2 |
| POCO X3 NFT | 7,26 | 1249 | 128 | 6 | 215 | 5160 | 2300 | 13 | 64 |
| MICROSOFT LUMIA 650 | 8,81 | 483 | 16 | 1 | 122 | 2000 | 1300 | 32 | 8 |

## Proste statystyki przedstawiające wykorzystane dane

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Ocena | Cena | Pamiec  wbudowana | RAM | Waga | Bateria | Procesor | Roznica kwartałów | Megapiksele |
| Średnia | 7,2 | 2884,6 | 272,7 | 9 | 189,7 | 4323,7 | 2596,96 | 7,74 | 56,38 |
| Odchylenie std | 1,06 | 2822,6 | 279,31 | 4,78 | 39,6 | 1168,1 | 890,88 | 9,4 | 40,48 |
| Min | 5 | 249 | 0,015625 | 0,016 | 79,6 | 1200 | 0 | 0 | 2 |
| Max | 8,81 | 11000 | 1000 | 16 | 246 | 6000 | 4200 | 32 | 200 |
| Q1 | 6,52 | 968,5 | 128 | 6 | 169 | 3692,5 | 2200 | 2 | 48 |
| Mediana | 7,28 | 1899 | 128 | 8 | 196 | 4520 | 2620 | 3 | 50 |
| Q3 | 8,09 | 3849,5 | 384 | 12 | 212,5 | 5001,5 | 3200 | 10 | 64 |

## Współczynnik zmienności

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Ocena | Cena | Pamięć  wbudowana | RAM | Waga | Bateria | Procesor | Różnica kwartałów | Megapiksele |
| Wsp zmienności | 14,5% | 95,7% | 100,18% | 51,9% | 20,41% | 26,42% | 33,55% | 118,8% | 70,27% |

## Macierz korelacji

A chart with different colored squares

Description automatically generated

Analizując macierz korelacji możemy być zaskoczeni faktem, że nie ma znaczącej korelacji między oceną telefonu a jego parametrami technicznymi. Ta obserwacja wynika z heterogeniczności analizowanych modeli telefonów, które pochodzą z różnych okresów. Parametry, które w obecnej erze technologicznej są uznawane za suboptymalne, w przeszłości były całkowicie konkurencyjne w stosunku do dostępnych na rynku standardów. Dodatkowo, zaobserwowano nieznaczącą, lecz pozytywną korelację między oceną a różnicą w kwartałach wprowadzenia modelu na rynek. Możliwym wyjaśnieniem tego zjawiska jest ewolucja oczekiwań użytkowników w stosunku do telefonów komórkowych, co może skutkować ich bardziej krytycznym podejściem do oceny nowszych modeli. Podobną korelację można zauważyć pomiędzy oceną a zegarem procesora. Co ciekawe, nie ma praktycznie żadnej korelacji pomiędzy oceną telefonu a innymi parametrami. Te fakty mogą wskazywać, że nie mała grupa konsumentów nie zwraca uwagi na procesor zainstalowany w telefonie, a następnie ocenia ten telefon między innymi na podstawie działania zainstalowanego w nim procesora.

Największa korelacja została zauważona dla przypadku pamięci wewnętrznej porównywanej z ceną. Korelacja 0,93 wynika prawdopodobnie z tego, że telefony z 512/1000 GB pamięci jest niewiele i wyróżniają się na rynku

Największa korelacja została zauważona dla przypadku pamięci wewnętrznej porównywanej z ceną. Korelacja 0,93 wynika prawdopodobnie z tego, że telefony z 512/1000 GB pamięci są jednocześnie modelami droższymi, ponieważ są modelami gamingowymi (np. ASUS ROG PHONE 7 także w wersji ultimate, nubia red magic 7 pro) lub markami premium (np. produkty Apple takie jak iPhone 15 pro max)

Występuje także istotna korelacja pomiędzy wagą a pojemnością baterii, której wartość oscyluje w okolicy 0,9. Podejrzewamy, że jest to spowodowane sporym udziałem baterii w wadze całego telefonu oraz powolnymi zmianami technologicznymi dotyczącymi baterii, co spowodowało, że telefony na przestrzeni kwartałów mają bardzo zbliżone stosunki masy baterii do energii przez nie magazynowanej.

Dodatkowo można zauważyć, iż wszystkie korelacje różnicy kwartałów poza ocenami są istotnie ujemne, co potwierdza postęp technologiczny w modelach telefonów dostępnych na rynku oraz rosnące ceny nowych modeli.

## Wartości odstające

### Identyfikacja

Po wstępnej analizie uwidaczniają się 3 wartości odstające. Są to modele:

1. MICROSOFT LUMIA 650
2. HTC 10
3. NOKIA 3310 2017 DUAL SIM

Telefony te są najstarszymi dostępnymi modelami przez nas zanotowanymi. Na tle konkurencji cechują się słabymi parametrami technicznymi, niskimi cenami i stosunkowo dobrymi opiniami.

Z tego też powodu dla dalszych badań usuwamy parametr wagi. Zdecydowaliśmy się na usunięcie właśnie tego parametru, gdyż wystarczająco dobrze tłumaczy pojemność baterii, oraz jest jednocześnie parametrem najmniej istotnym z omawianej dwójki. Dane natomiast będą różnić się w zależności od wykonywanych badań. Ze względu na istotność współczynnika zmienności dla analizy skupień, wyłącznie dla owego badania nie będziemy uwzględniali wartości odstających. Dzięki temu zabiegowi współczynnik zmienności istotnie się zmieni.

### Efekt dokonanych zmian

Obraz zawierający tekst, zrzut ekranu, numer

Opis wygenerowany automatycznie

Współczynnik zmienności:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Ocena | Cena | Pamięć  wbudowana | RAM | Waga | Bateria | Procesor | Różnica kwartałów | Megapiksele |
| Wsp zmienności | 13,56% | 85,42% | 87,66 | 38,84% | 13,61% | 15,77% | 22,43% | 88,47% | 58,45% |

### Dodatkowe spostrzeżenia

Eliminacja najstarszych modeli