

Analiza rynku kart graficznych na podstawie danych Steam Hardware Survey

Autorzy

Sebastian Bek s223502

Adam Janus s223530

Aliaksei Kashynski s217616

Jakub Mierzwa s217562

Streszczenie

W niniejszej pracy dokonano analizy rynku kart graficznych na podstawie danych pochodzących z ankiety Steam Hardware Survey, obejmujących popularne modele producentów NVIDIA, AMD oraz Intel. W pracy przedstawiono wstępne oczyszczenie danych, statystyki opisowe i wnioski dotyczące rozkładu parametrów takich jak pojemność VRAM, pobór mocy (TGP), wydajność karty (GPU Score), sugerowana cena detaliczna (MSRP) oraz udział w rynku według Steama. Następnie zrealizowano wielokryterialne porządkowanie kart graficznych przy wykorzystaniu metod normalizacji i agregacji. Wynikiem końcowym jest porównanie uzyskanych rankingów oraz omówienie, które modele kart graficznych zajmują najwyższe miejsca przy różnych założeniach analizy. Projekt podsumowano oceną stopnia realizacji założonych celów.

Słowa kluczowe

Karty graficzne, analiza rynku, suma ważona, iloczyn ważony, TOPSIS, porządkowanie

Wprowadzenie

Rynek kart graficznych jest dynamiczny i zróżnicowany, a wybór odpowiedniego modelu zależy od wielu czynników, takich jak wydajność, cena czy efektywność energetyczna. Steam Hardware Survey to jedno z największych źródeł danych o sprzęcie komputerowym używanym przez graczy. W niniejszej pracy wykorzystano dane z tej ankiety do analizy kart graficznych pod względem ich parametrów technicznych i popularności wśród użytkowników. Zastosowanie metod wielokryterialnych w analizie sprzętu komputerowego opisywane jest w wielu publikacjach naukowych. Przykładowo, Kisielińska i Stanko (2011) wskazują, że techniki porządkowania obiektów, takie jak suma ważona, iloczyn ważony czy TOPSIS, są skuteczne przy ocenie wariantów różniących się wieloma cechami – zwłaszcza jeśli część z nich jest pożądana, a inne niepożądane. W

badaniach nad rynkiem technologii IT metody te wykorzystywane są do oceny efektywności, opłacalności oraz popularności różnych rozwiązań sprzętowych (Kukuła, 2000).

Przedmiot badania

Cel i zakres badania

Celem analizy jest porównanie i uszeregowanie kart graficznych dostępnych na rynku na podstawie ich parametrów technicznych oraz popularności wśród użytkowników platformy Steam, z wykorzystaniem metod wielokryterialnego porządkowania (suma ważona, iloczyn ważony, TOPSIS). Na podstawie uzyskanych rankingów dokonano również analizy pozycji poszczególnych producentów (NVIDIA, AMD, Intel), co pozwala na ocenę ich konkurencyjności w różnych aspektach takich jak wydajność, efektywność energetyczna, cena i udział rynkowy oraz ocenę zależności między cechami technicznymi kart a ich popularnością.

Przegląd literatury

Zmienne

- Ilość pamięci VRAM (w GB),
- Pobór mocy TGP (w watach)
- Wydajność karty (GPU Score)
- Sugerowana cena detaliczna (MSRP w dolarach)
- Udział procentowy w rynku według ankiety Steam (Steam February %)

- Producent, zmienna jakościowa

Zmienna **VRAM**, **GPU Score** oraz **Udział procentowy w rynku według ankiety Steam** stanowią stymulanty, ponieważ wyższe wartości są pożądane. Z kolei **TGP** oraz **MSRP** traktowane są jako destymulanty, ponieważ karty o niższym poborze mocy i cenie są preferowane przez konsumentów.

Wstępna analiza danych

Filtracja modeli kart mobilnych

Pierwszym krokiem było usunięcie kart graficznych przeznaczonych do laptopów, ponieważ głównym celem analizy jest rynek kart dla komputerów stacjonarnych. Karty mobilne zidentyfikowano na podstawie występowania określonych fraz w nazwie modelu.

W wyniku tej operacji usunięto 11 modeli kart mobilnych, pozostawiając 58 modeli kart stacjonarnych do dalszej analizy.

Normalizacja udziału procentowego

Zauważono, że suma udziałów procentowych wszystkich kart w zbiorze danych wynosi 76,08%, a nie 100%. Wynika to z faktu, że w oryginalnym badaniu Steam uwzględniono również inne karty graficzne, które nie zostały ujęte w naszym zbiorze danych.

Aby uzyskać bardziej intuicyjną interpretację udziałów rynkowych, przeprowadzono normalizację procentów, tak aby sumowały się do 100%. Zastosowano następującą formułę:

$$\text{nowy_procent} = (\text{obecny_procent} / \text{suma_wszystkich_procentów}) * 100$$

Normalizacja ta pozwala na lepsze porównanie popularności poszczególnych modeli kart graficznych w analizowanym zbiorze danych.

Statystyki opisowe

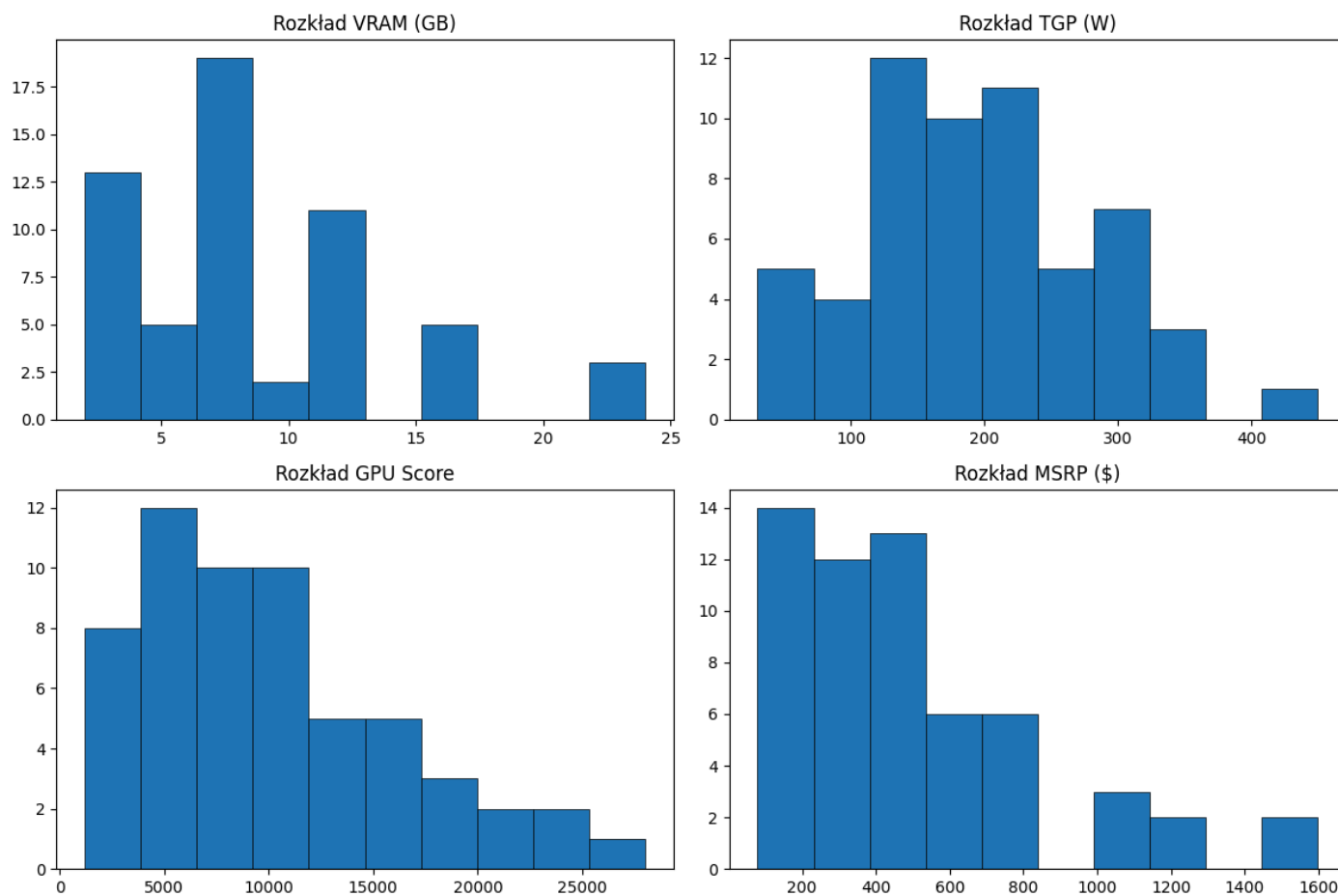
Podstawowe charakterystyki przyjętych cech diagnostycznych

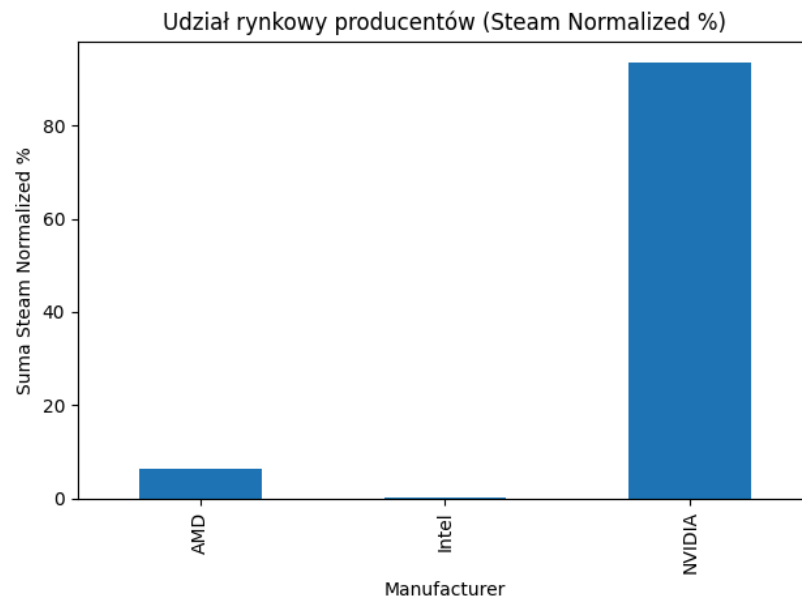
Zmienna	Min	Max	Średnia	Mediana	Odch. std.	Skośność
VRAM (GB)	2	24	9	8	5,16	1,18
TGP (W)	30	450	192,19	180	89,9	0,41
GPU Score	1200	28000	10079,31	9100	6218,7	0,79
MSRP (USD)	79	1599	487,28	399	336,9	1,445
Udział w rynku (%)	0,20	11,24	1,72	0,71	2,398	2,388

Podstawowa wizualizacja

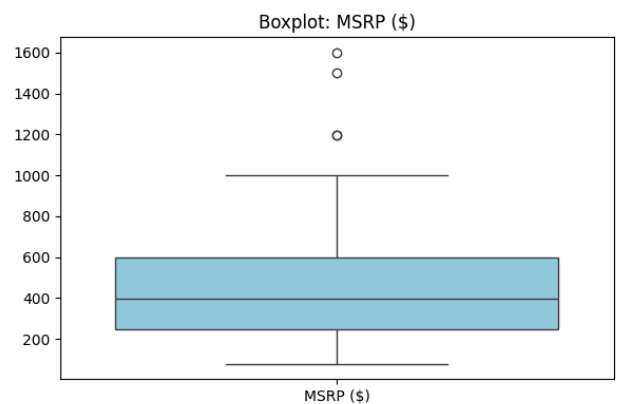
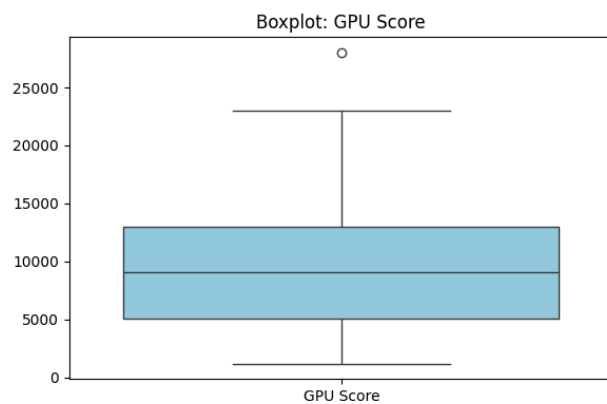
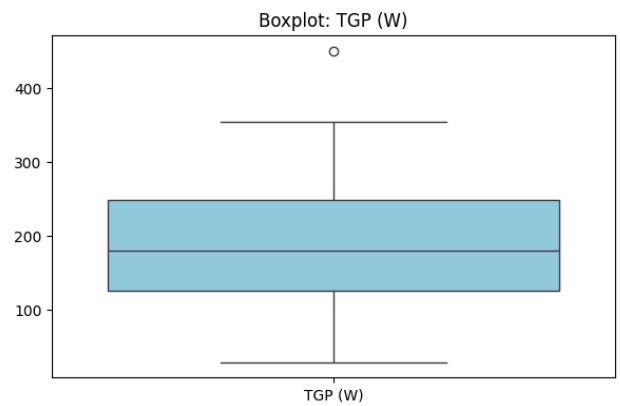
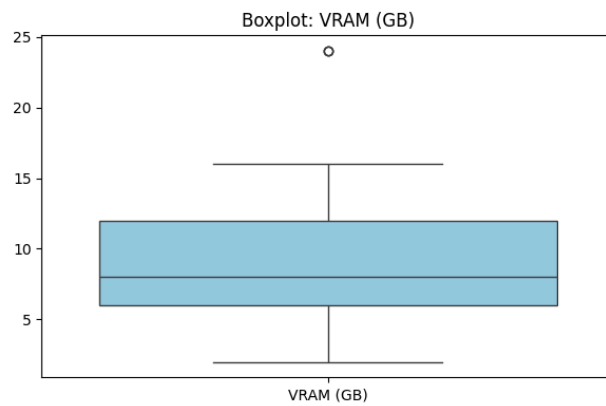
Do analizy danych zastosowano wykresy **boxplot** i **histogramy**, które obrazują rozkład poszczególnych zmiennych, co pozwala na lepszą interpretację wyników.

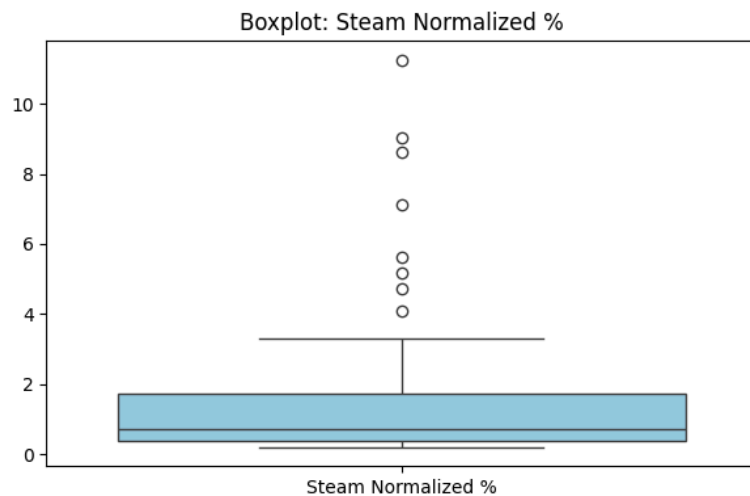
Histogramy przyjętych cech diagnostycznych



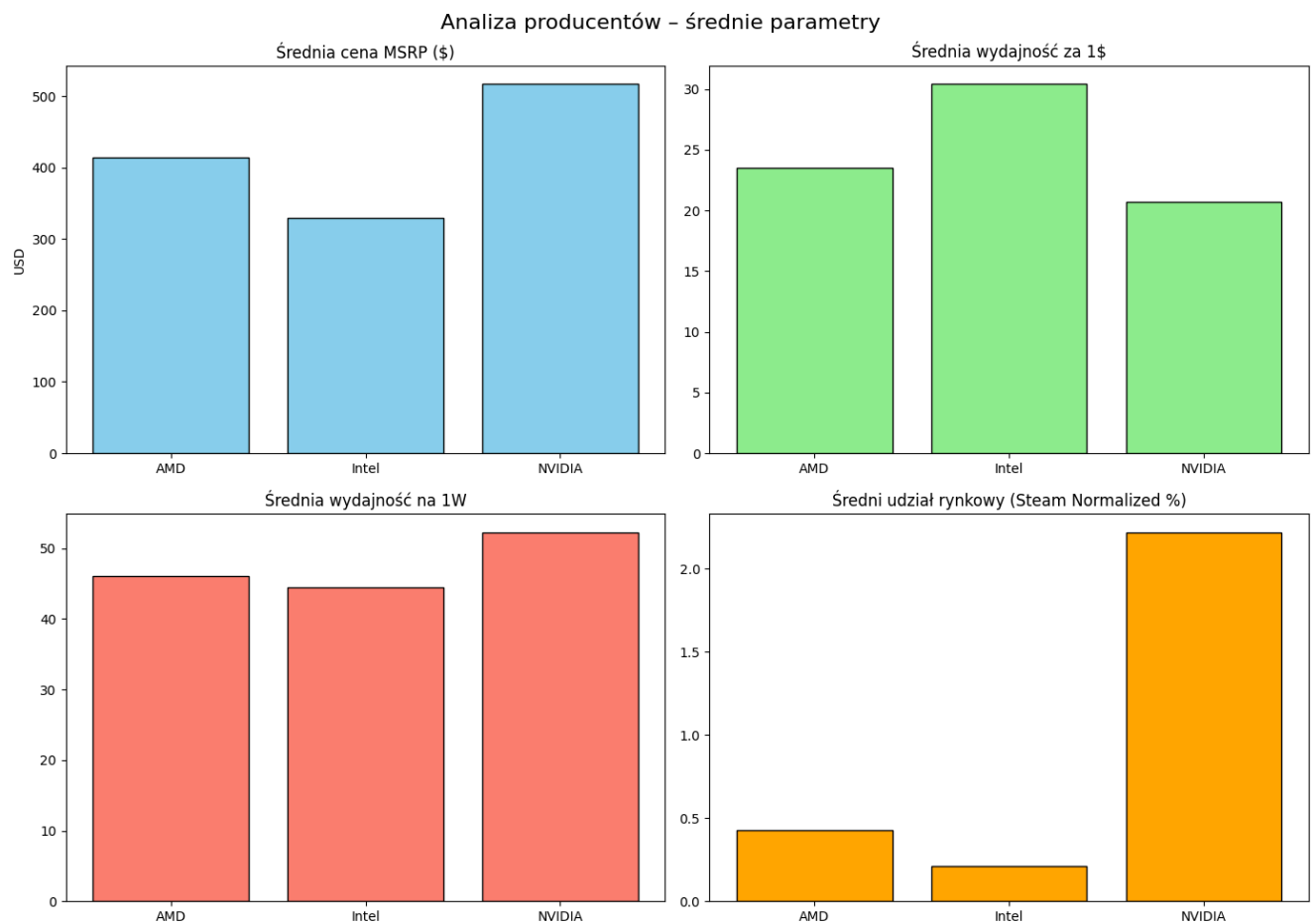


Wykresy boxplot przyjętych cech diagnostycznych

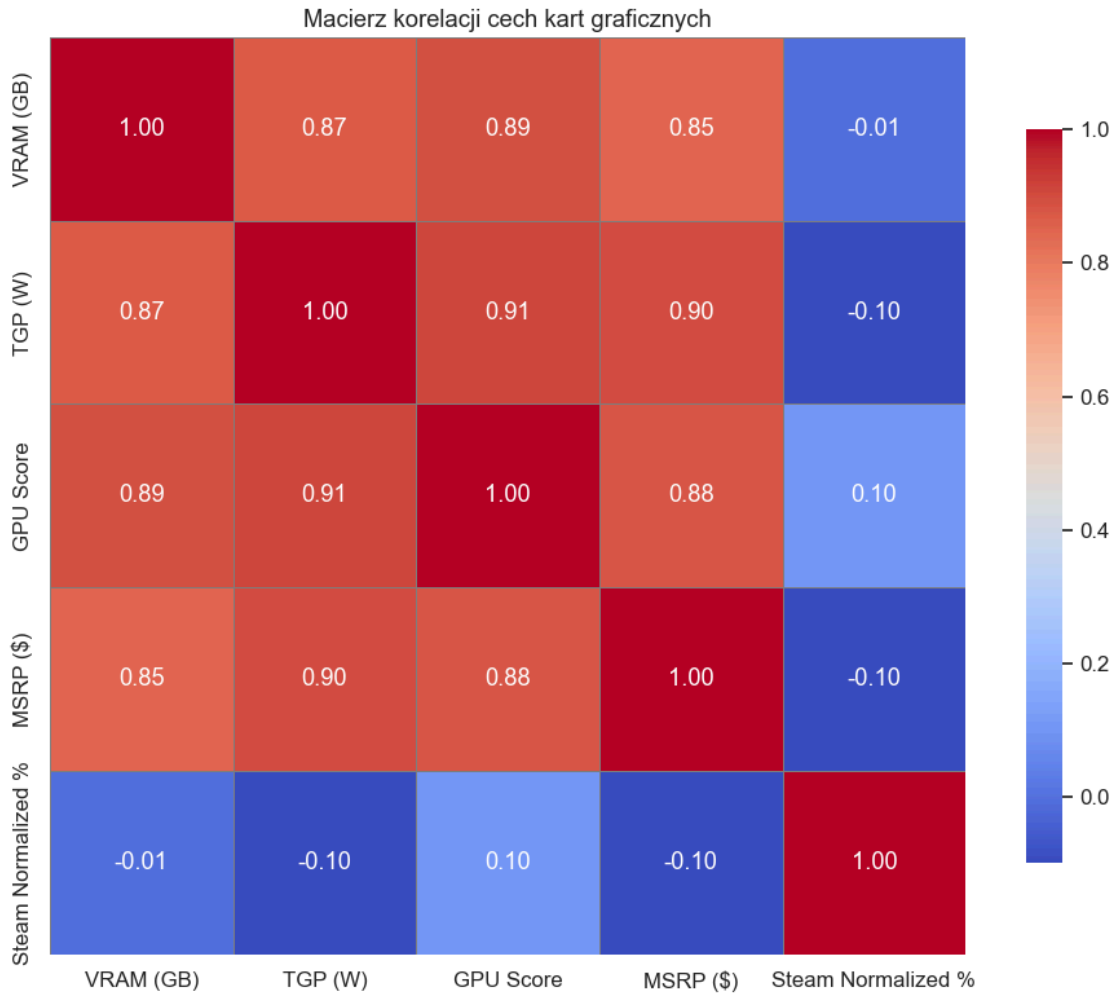




Dodatkowo przedstawiono wartości następujących parametrów: **średnia cena MRSP**, **średnia wydajność za 1\$**, **średnia wydajność na 1W** oraz **średni udział rynkowy** według poszczególnych producentów.



Ponadto wykonano analizę korelacji między parametrami technicznymi kart i udziałem w rynku. Wyniki analizy przedstawiono w tabeli:



Braki danych i ich obsługa

W badanym zbiorze danych nie stwierdzono braków. Wszystkie dostępne informacje zostały uwzględnione w analizie.

Obserwacje odstające

Obserwowano kilka danych odstających w zakresie poboru mocy i wydajności GPU. Zostały one uwzględnione w analizie, jako że są istotne w kontekście różnorodności dostępnych modeli.

Opis metod

Suma ważona

Metoda ta polega na agregacji wartości znormalizowanych cech dla każdego obiektu (karty graficznej) z uwzględnieniem przypisanych wag. Metoda ta jest intuicyjna i szeroko stosowana, choć podatna na wpływ jednostek miary – stąd konieczna była wcześniejsza normalizacja danych. Wysoki wynik oznacza większą „atrakcyjność” karty graficznej według zadanych kryteriów. Metoda ta, opisana szczegółowo przez Kukułę (2000), jest jedną z podstawowych technik agregacji ocen i często stosowana w analizach rynku, finansach i zarządzaniu projektami. W przypadku analizy kart graficznych pozwala ona w przejrzysty sposób uwzględnić zarówno wydajność, jak i cenę oraz inne czynniki.

Iloczyn ważony

W tej metodzie agregacja opiera się na przemnożeniu znormalizowanych wartości zmiennych, każda podniesiona do potęgi odpowiadającej wadze. Metoda ta charakteryzuje się większą odpornością na skrajne wartości (outliery) i zachowuje proporcje między ocenianymi obiektami. W praktyce wykorzystywana jest w analizach, gdzie relatywne różnice między zmiennymi mają duże znaczenie. Według Kisielińskiej i Stanko (2011), metoda iloczynowa jest bardziej odporna na duże rozrzuty w wartościach zmiennych i lepiej odzwierciedla relacje proporcjonalne między obiektami. Jest szczególnie ceniona w analizach, gdzie nie chodzi wyłącznie o sumę punktów, ale o zachowanie pewnych proporcji pomiędzy ocenianymi cechami.

TOPSIS

TOPSIS to metoda porządkowania obiektów przez porównanie ich odległości od tzw. rozwiązania idealnego i antyidealnego. Dla każdego obiektu wyznacza się odległość euklidesową od punktu idealnego (najlepsze możliwe wartości cech) oraz punktu antyidealnego (najgorsze możliwe wartości cech). Wartość końcowa (miara bliskości) obliczana jest według wzoru przedstawionego w następnym akapicie. Im wyższa wartość C_i , tym bliżej obiekt znajduje się rozwiązania idealnego, a zatem uznawany jest za bardziej korzystny. Metoda TOPSIS dobrze sprawdza się w przypadku wielu zmiennych o różnych kierunkach preferencji (stymulanty i destymulanty), ponieważ uwzględnia jednocześnie najlepszy i najgorszy możliwy scenariusz. TOPSIS został zaproponowany przez Hwang i Yoon (1981) i jest jedną z najbardziej znanych metod wspomagania decyzji

wielokryterialnych. Jego skuteczność została potwierdzona w licznych badaniach, zarówno w obszarze ekonomii, inżynierii, jak i analizy rynku (Kisielińska i Stanko, 2011).

Wzory

1. Suma prosta ważona

$$R_i = \sum_{j=1}^m w_j \cdot z_{ij}$$

gdzie:

- m to liczba cech,
- w_j to waga cechy j ,
- z_{ij} to znormalizowana wartość cechy j dla obiektu i .

2. Iloczyn ważony

$$R_i = \prod_{j=1}^m (z_{ij})^{w_j}$$

gdzie:

- m to liczba cech,
- w_j to waga cechy j ,
- z_{ij} to znormalizowana wartość cechy j dla obiektu i .

3. TOPSIS

$$C_i = \frac{D_i^-}{D_i^+ + D_i^-}$$

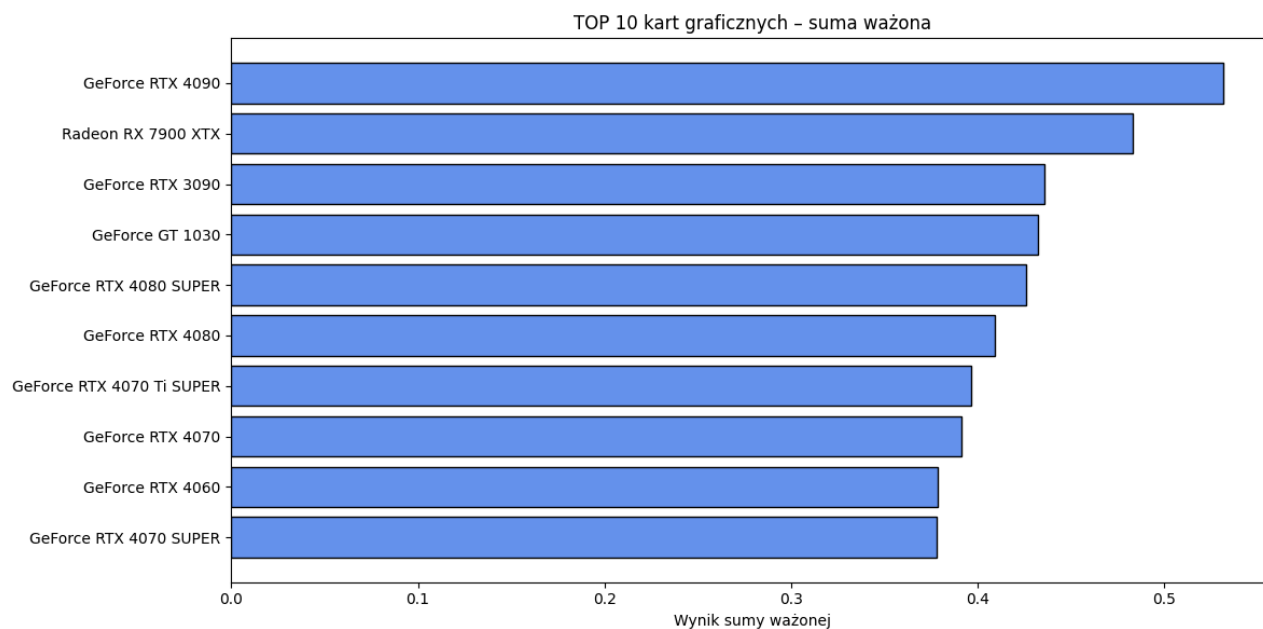
gdzie :

- C_i – wskaźnik bliskości obiektu i do rozwiązania idealnego ($0 < C_i < 1$),
- D_i^+ – odległość od rozwiązania idealnego,
- D_i^- – odległość od rozwiązania antyidealnego.

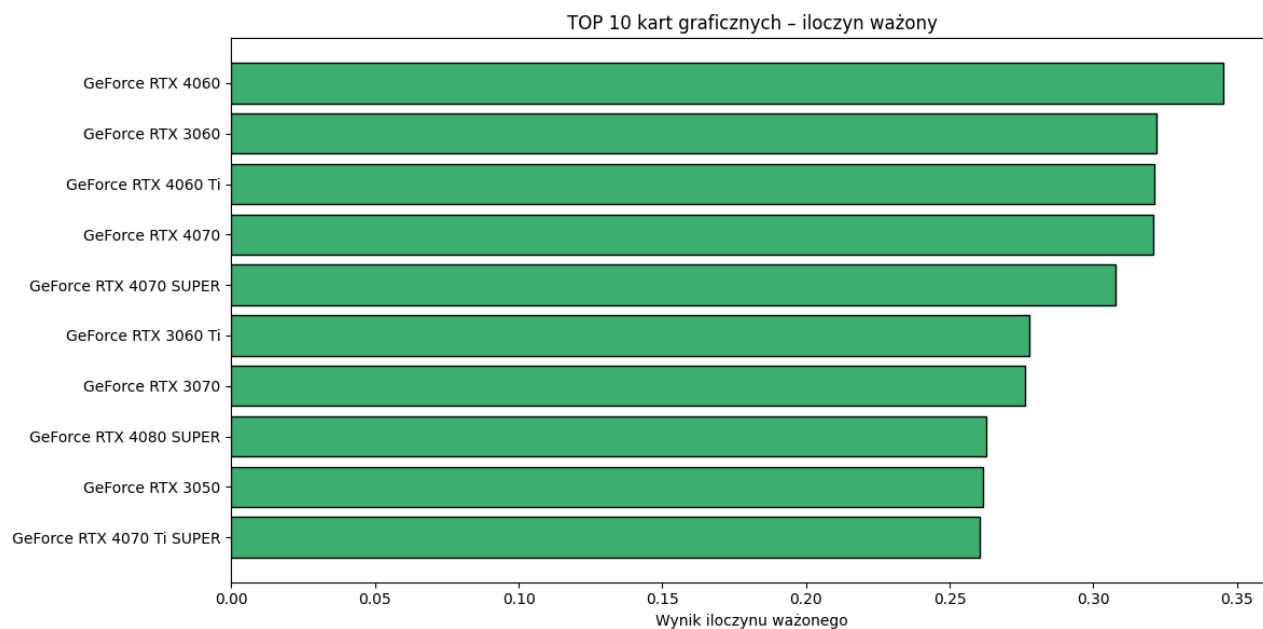
Rezultaty

Poniżej przedstawiono wyniki porządkowania (top 10) dla każdej z metod.

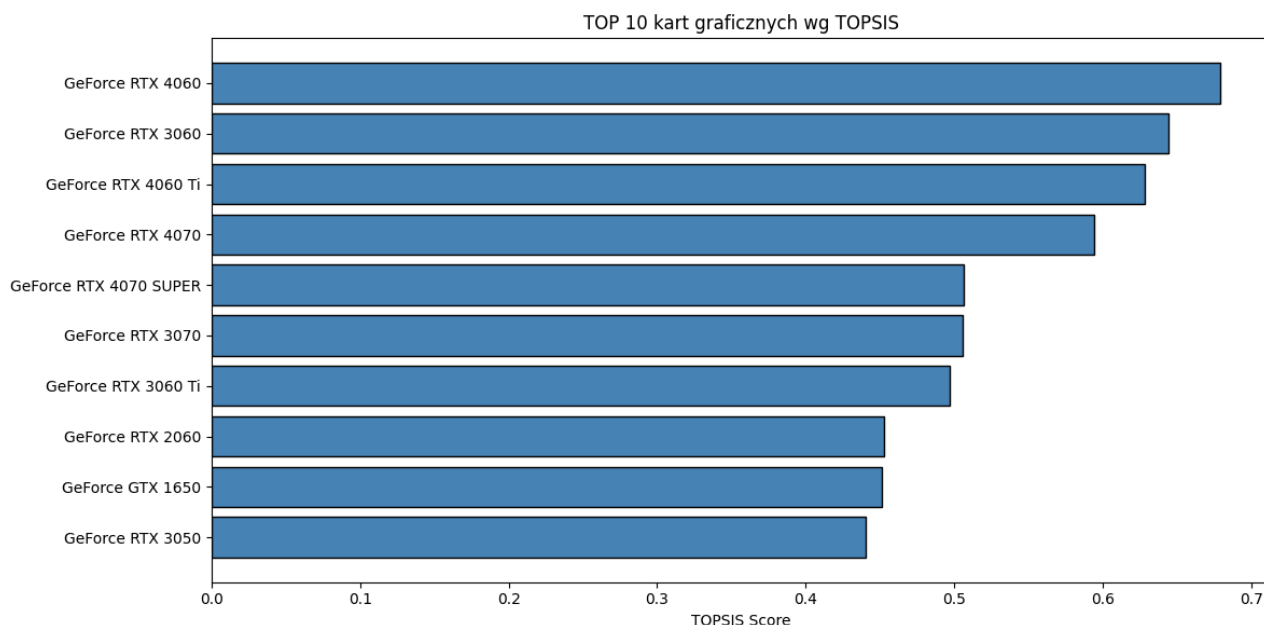
Suma ważona



Iloczyn ważony



TOPSIS



W wynikach analizy zauważalna jest wyraźna dominacja kart graficznych firmy NVIDIA – we wszystkich trzech metodach porządkowania (suma ważona, iloczyn ważony, TOPSIS) to właśnie ich modele zajmują całe pierwsze dziesiątki rankingów (oprócz jednego modelu dla sumy ważonej). Może to wynikać z połączenia wysokiej wydajności tych kart oraz bardzo dużego udziału w rynku – aż 95% według danych Steam. Ponieważ udział w rynku został w analizie potraktowany jako stymulanta i uwzględniony przy ocenie ogólnej, NVIDIA zyskała tu naturalną przewagę.

Ciekawe jest jednak to, że pod względem stosunku wydajności do ceny oraz efektywności energetycznej, różnice między NVIDIA a AMD nie są aż tak duże. Średnia cena kart NVIDIA jest wprawdzie wyższa (ok. 500 USD vs 400 USD dla AMD), ale wydajność na 1W i wydajność za 1 USD wypadają dość podobnie. Może to oznaczać, że przewaga NVIDIA niekoniecznie wynika wyłącznie z parametrów technicznych, ale też z marketingu, popularności czy większej dostępności.

Warto też zwrócić uwagę na skośność w danych – większość zmiennych (np. GPU Score, MSRP, VRAM) ma rozkłady z długim „ogonem” w prawo. To oznacza, że tylko kilka modeli ma naprawdę wysokie wartości, i to właśnie one – w większości od NVIDIA – mocno wpływają na końcowy wynik rankingu. To dodatkowo wzmacnia efekt dominacji tej marki w analizie.

Jednocześnie analiza macierzy korelacji pokazuje, że udział rynkowy kart nie jest silnie powiązany z ich parametrami technicznymi. Najmocniejsze modele, choć dominują w rankingach, są stosunkowo rzadko używane. Największą popularnością cieszą się zaś karty ze średniego segmentu, co sugeruje, że użytkownicy w praktyce częściej kierują się kwestiami finansowymi niż maksymalnymi osiągnięciami.

Podsumowanie

W pracy przeanalizowano karty graficzne na podstawie danych ze Steam Hardware Survey, wykorzystując metody wielokryterialnego porządkowania (suma ważona, iloczyn ważony, TOPSIS). Wyniki pokazały wyraźną dominację modeli NVIDIA, co wynika głównie z ich wysokiej wydajności i dużego udziału w rynku.

Choć pod względem efektywności energetycznej i stosunku wydajności do ceny różnice między producentami nie są duże, to silna pozycja NVIDIA wynika też z popularności i obecności w danych. Skośność rozkładów sprzyjała topowym modelom, dodatkowo wzmacniając ten efekt.

Uzyskane wyniki stanowią dobrą bazę do dalszej analizy, np. grupowania kart czy porównania strategii producentów. Zastosowanie trzech różnych metod porządkowania pozwoliło nie tylko porównać efekty analizy, ale także wykazać, że wyniki są spójne niezależnie od sposobu agregacji informacji. Taki wniosek potwierdza również literatura – jak zauważa Kukuła (2000) zastosowanie różnych podejść pozwala na lepszą ocenę obiektów i ograniczenie ryzyka błędnej interpretacji danych.

Załączniki

1. Repozytorium GitHub

Wszystkie pliki wykorzystane w niniejszym projekcie – w tym dane, skrypty w języku Python (do analiz, normalizacji, porządkowania oraz wizualizacji danych), wykresy oraz plik z końcowym raportem – zostały umieszczone w repozytorium GitHub, dostępnym pod adresem:

<https://github.com/AdamN3wTS/mad-steam>

Repozytorium zawiera:

- plik `Data.csv` z danymi wejściowymi,
- skrypty `.py` do analizy danych (w tym: suma ważona, iloczyn ważony, TOPSIS),
- folder `img/` z wygenerowanymi wykresami,
- gotowy raport `.docx` oraz plik `.pdf`,
- plik `requirements.txt` z listą wymaganych bibliotek.

Repozytorium ma na celu zapewnienie pełnej przejrzystości i umożliwienie łatwego odtworzenia wyników analizy.

Bibliografia

1. Kukuła, K. (2000). Metoda unitaryzacji zerowanej. Warszawa: Wydawnictwo Naukowe PWN.
2. Kisielińska, J., & Stanko, S. (2011). Wielokryterialne metody wspomaganie decyzji. Kraków: Uniwersytet Ekonomiczny w Krakowie.
3. Hwang, C. L., & Yoon, K. (1981). Multiple Attribute Decision Making. Berlin: Springer-Verlag.
4. Steam Hardware Survey (luty 2024). <https://store.steampowered.com/hwsurvey/>