

# Porównanie

Kacper Bagiński

2024-04-02

Wykres pochodzi z postu na Reddicie z dn. 26.03.2024. i przedstawia ile procent użytkowników Steama ma ile rdzeni w procesorze.

Udało mi się odnaleźć dane, których użył autor. Pochodzą one z Ankiety Steama z kwietnia 2022. Nie wiem, czemu autor akurat ich użył. Nie wiem też dla jakich dokładnie systemów autor wybierał swoje dane. Ja ustaliłem, że chodziło pewnie o Windowsa i dane na wykresie wyglądają podobnie.

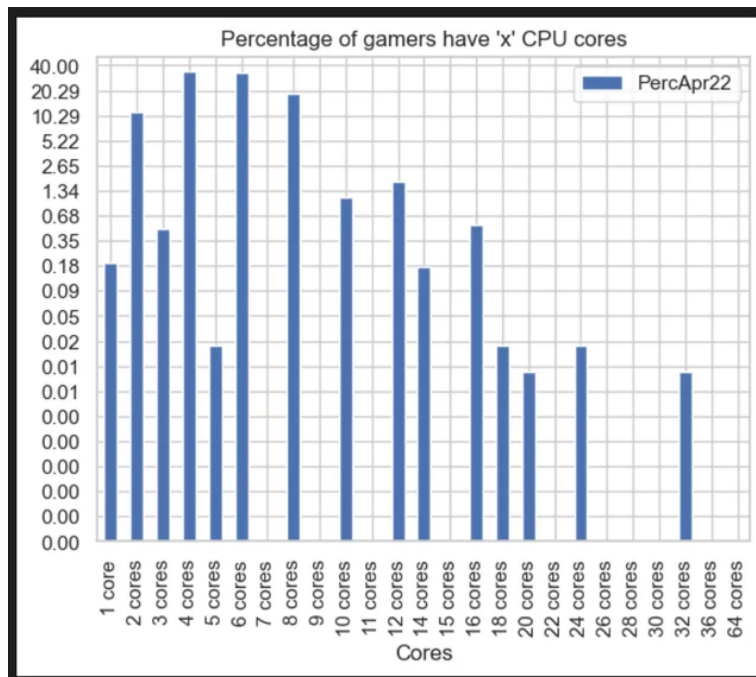


Figure 1: Oryginalny Wykres

- oczywistym problemem jest to, że autor wykorzystał skalę logarytmiczną, żeby porównać wartości procentowe. Patrząc na słupki nie da się stwierdzić ile razy więcej jest jednego rodzaju procesorów od innego. Autor próbował na siłę, porównać dane o bardzo różnych wielkościach.
- Ponadto przez zastosowanie skali logarytmicznej i zaokrąglania, na osi y, znalazło się 6-krotnie 0.00, co już kompletnie odbiera sens tym niższym słupkom. Gdyż nawet wspomagając się y nie możemy porównać ich wartości.
- Tło z równoodległymi prostymi implikuje raczej skalę liniową. I trzeba się chwilę zastanowić, żeby zauważyć ten problem z y.

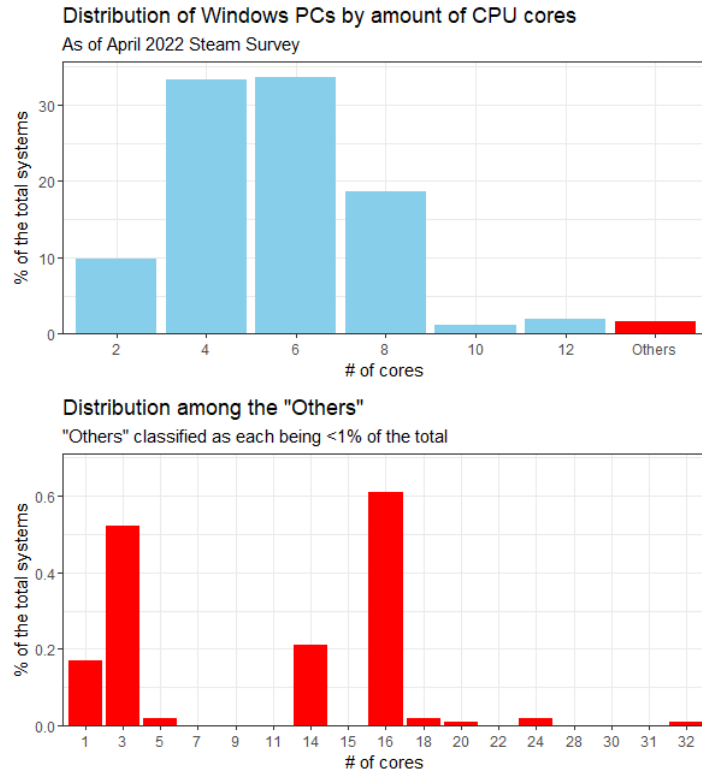


Figure 2: Mój wykres

Zastąpiłem skalę logarymiczną, intuicyjniejszą skalą liniową. Wartości poniżej <1% zsumowałem i zaznaczyłem kolorem jako “Others”. I porównanie tych malutkich wartości umieściłem na oddzielnym wykresie pod spodem

Naprawia to większość opisanych problemów.

Co do skali x, to postanowiłem zostawić sortowanie względem ilości rdzeni. W komentarzach pod oryginalnym wykresem były mieszane o tym opinie. Ale ja myślę, że jest bardziej intuicyjnie sortować względem ilości rdzeni niż wysokościami słupków. Bo jak osoba chce specyficzną ilość rdzeni sprawdzić, to w miarę szybko da się dojść, gdzie jest jej słupek. Też się akurat ładnie złożyło, że co dwa rdzenie są te pierwsze słupki.

Zostawiłem też puste miejsca. Ponieważ nie są to “prawdziwie puste” miejsca. Prawdziwie pustym miejscem jest np. 29 rdzeni, bo Steam nie uwzględnił takiej opcji w ankiecie. Natomiast np. opcja 31 rdzeni istniała w ankiecie, ale zyskała tak mało odpowiedzi, że w ramce danych została zaokrąglona do 0 procent wszystkich odpowiedzi. A więc istnienie tych pustych miejsc na wykresie daje dodatkową informację, o tym jakie odpowiedzi w ogóle istniały.

---

Link do postu: [https://www.reddit.com/r/pcmasterrace/comments/1borv2t/which\\_one\\_of\\_you\\_folks\\_has\\_a\\_5\\_core\\_cpu\\_never/](https://www.reddit.com/r/pcmasterrace/comments/1borv2t/which_one_of_you_folks_has_a_5_core_cpu_never/)

i do danych: [https://raw.githubusercontent.com/jdegene/steamHWSurvey/master/shs\\_platform.csv](https://raw.githubusercontent.com/jdegene/steamHWSurvey/master/shs_platform.csv)