

# Najlepsze albumy muzyczne

wg użytkowników RateYourMusic.com

Bartosz Łuksza, Rafał Głodek

## Wprowadzenie

Muzyka towarzyszy człowiekowi od tysięcy lat. Zawsze stanowiła nieodłączną część naszej kultury. Niestety ograniczenia technologiczne przez długi czas nie pozwalały artystom utrwalić swoich dzieł. Fonografia narodziła się w XIX wieku, a jej największy rozkwit przypada na drugą połowę wieku XX. Z tego względu najstarsze oryginalne dzieła muzyczne, do których mamy obecnie dostęp, pochodzą poprzedniego stulecia. W ostatnich latach rynek muzyczny przeżywa niebywały rozkwit. Każdego roku miliardy słuchaczy na całym świecie, przesłuchuje miliony nowych albumów, generując przychody rzędu dziesiątek miliardów dolarów ze sprzedaży nagrań. Rynek muzyczny jest jednak ściśle powiązany z wieloma innymi gałęziami biznesu, takimi jak: film, moda, czy technologie cyfrowe. Szacuje się, że każdego dnia na serwisy streamingowe trafia nawet 120 000 utworów! W tej sytuacji można pokusić się o stwierdzenie, że obecny przemysł muzyczny jest wręcz “przeladowany” muzyką. Warto zadać sobie pytanie, czy za ilość idzie również jakość?

W naszej pracy zajrzemy wгłęb współczesnej historii muzyki i przeanalizujemy bazę pięciu tysięcy najlepiej ocenianych albumów muzycznych przez użytkowników RateYourMusic.com - największego portalu do oceniania muzyki w internecie. Dane pochodzą z 12 grudnia 2021 r. i zostały pobrane z serwisu kaggle.com. Wyodrębniliśmy z nich następujące zmienne:

1. **Album** - nazwa albumu
  - Zawiera 4928 unikalne wartości
2. **Artist Name** - artysta (imię i nazwisko lub pseudonim artystyczny)
  - Zawiera 2787 unikalne wartości
  - 25 rekordów to *Various Artists*, czyli różni artyści, których jednak nie możemy wyodrębnić, więc pomijamy te wartości
3. **Release Date** - dokładna data wydania albumu (dzień/miesiąc/rok), wyodrębniliśmy z niej dwie zmienne:
  - a) **Year** - rok wydania albumu

- najmniejsza wartość: 1947
- największa wartość: 2021
- średnia arytmetyczna: 1987,46
- mediana: 1988
- wariancja: 253,23

b) **Month** - miesiąc wydania albumu

4. **Genres** - gatunki (lub gatunek), do których należy album

- Niekiedy trudno jest ustalić, do jakiego gatunku należy album. Wówczas określany jest mianem międzygatunkowego i klasyfikuje się go jako przynależnego do każdego z wymienionych gatunków.
- Możliwe wartości to: "Rock", "Hip Hop", "Pop", "Jazz", "Soul", "Dance", "Techno", "Punk", "Metal", "Folk"

5. **Descriptors** - krótki opis albumu

- Opis zawiera kilka przymiotników najlepiej oddających charakter albumu, np. "melancholic, anxious, futuristic, alienation, existential, male vocals, atmospheric, lonely, cold, introspective"
- Opisy będą potrzebne, by sprawdzić jak oceniane są albumy w zależności od nastroju, jaki wywołują w słuchaczu

6. **Average Rating** - średnia ocen użytkowników

- Na stronie RateYourMusic.com użytkownicy mogą wystawiać albumom oceny w skali od 0 do 5, z uwzględnieniem "połówek"
- Średnie oceny, które są brane pod uwagę w tym spisie uwzględniają także wagi ocen - oceny użytkowników wykazujących się dużą aktywnością i doświadczeniem mają wyższą wagę niż tych, którzy oceniają muzykę sporadycznie
- najmniejsza wartość: 3,52
- największa wartość: 4,34
- średnia arytmetyczna: 3,771
- mediana: 3,75
- wariancja: 0,0098

7. **Number of Ratings** - liczba ocen użytkowników

- najmniejsza wartość: 260
- największa wartość: 70 400

- średnia arytmetyczna: 4084.511
- mediana: 1820
- wariancja: 36016085

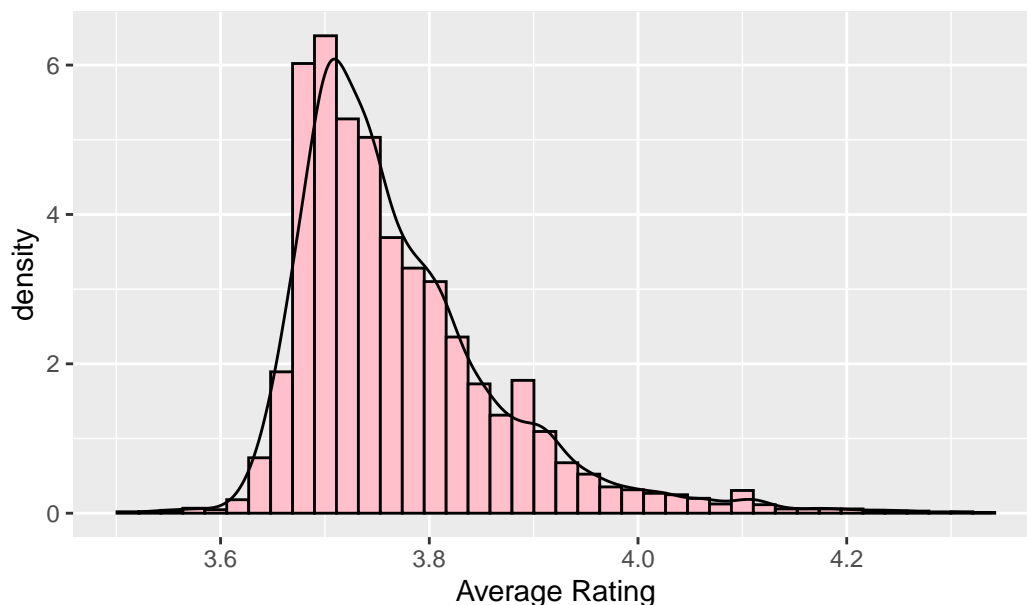
8. **Number of Reviews** - liczba recenzji użytkowników

- najmniejsza wartość: 0
- największa wartość: 1 549
- średnia arytmetyczna: 71.4492
- mediana: 34
- wariancja: 11766.56

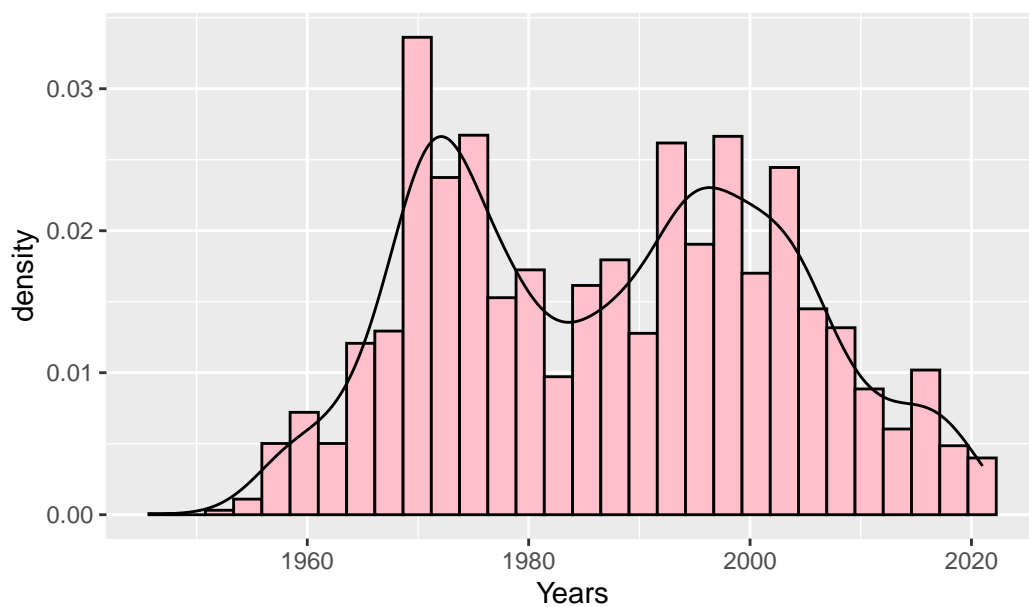
Dogłębna analiza pozwoli nam znaleźć korelacje między różnymi zmiennymi ujętymi w zestawieniu i wyciągnąć nieoczywiste wnioski. W ten sposób nie tylko dowiemy się, jak muzyka rozwijała się w ubiegłych dekadach, ale także nakreślimy ścieżkę jej dalszego rozwoju.

W jakich latach powstawało najwięcej “dobrych” albumów? Czy istnieje korelacja między średnią oceną użytkowników a datą wydania dzieła? Jakie są średnie ocen dla różnych gatunków muzycznych? Którzy artyści mogą się poszczycić najlepiej ocenianą dyskografią? Na te i wiele innych pytań odpowiemy w naszej pracy.

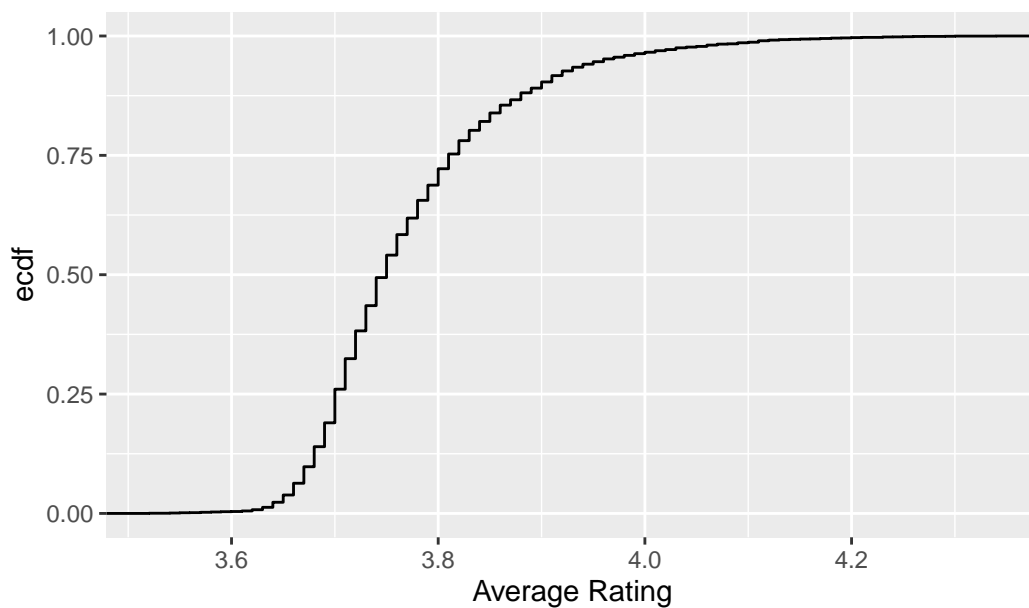
Histogram gestosci dla srednich ocen

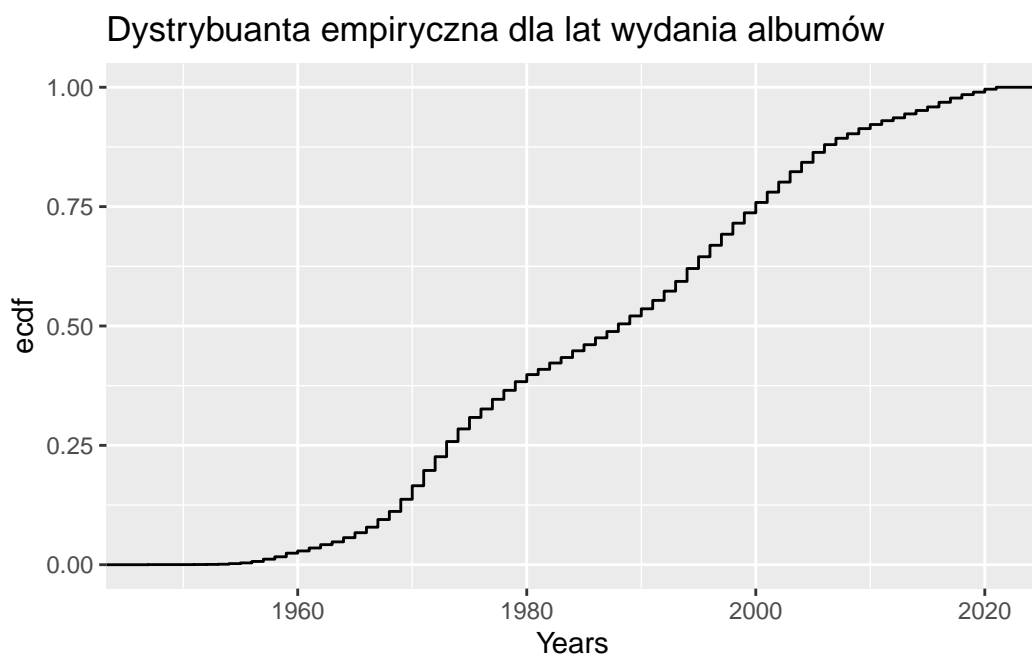


Histogram gestosci dla srednich ocen

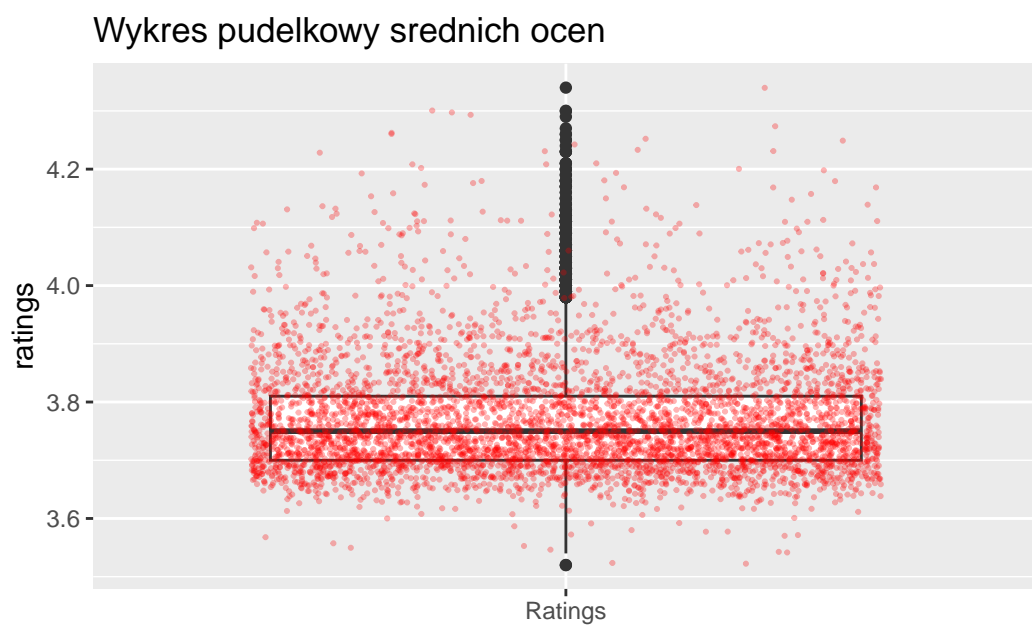


Dystrybuanta empiryczna dla lat wydania albumów

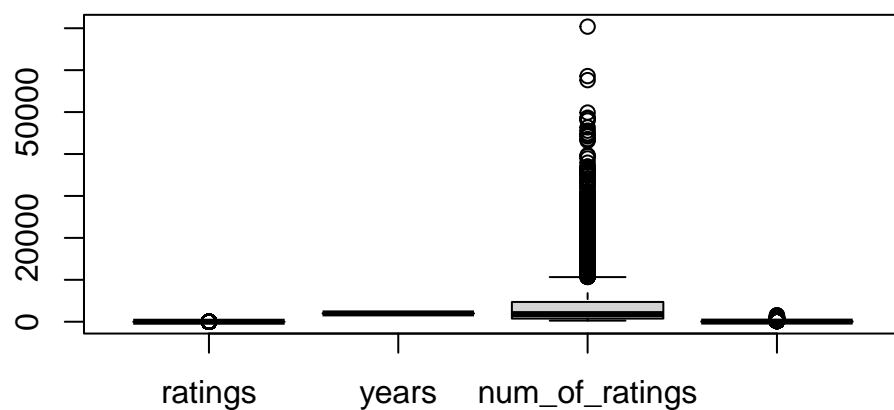
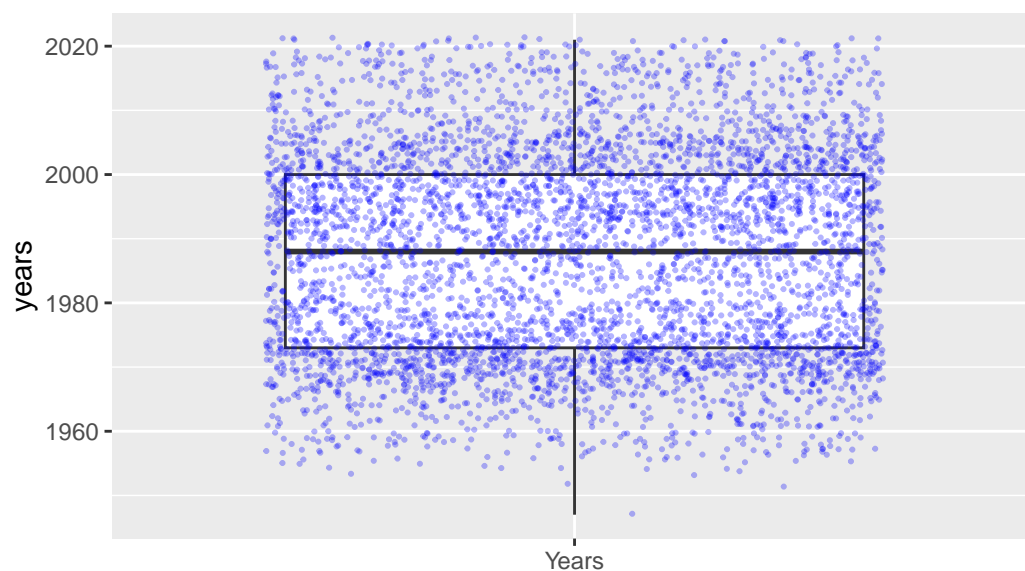




Ładowanie wymaganego pakietu: viridisLite



Wykres pudełkowy dla lat wydania albumów



	Ratings	Years	Number of Ratings	Number of Reviews
Min	3.54	1947	260.0	0.0
First Quartile	3.70	1973	752.5	14.0
Median	3.75	1988	1820.0	34.0
Third Quartile	3.81	2000	4720.5	82.5
Maximum	3.97	2021	10649.0	185.0

```
lm_ratings_and_years = lm(ratings~years)
summary(lm_ratings_and_years)
```

Call:

```
lm(formula = ratings ~ years)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.25179	-0.06707	-0.02552	0.04131	0.54285

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	5.871e+00	1.724e-01	34.05	<2e-16 ***
years	-1.057e-03	8.676e-05	-12.18	<2e-16 ***

---

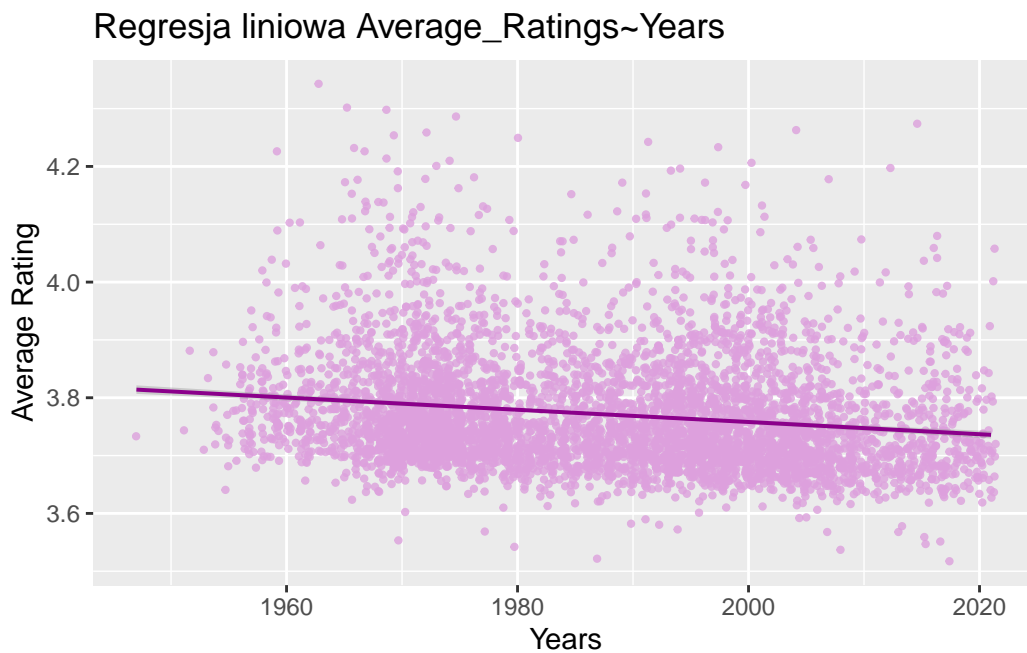
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.09762 on 4998 degrees of freedom

Multiple R-squared: 0.02882, Adjusted R-squared: 0.02863

F-statistic: 148.3 on 1 and 4998 DF, p-value: < 2.2e-16

```
`geom_smooth()` using formula = 'y ~ x'
```



```
library(ggplot2)
```

```
ggplot() +  
  geom_point(aes(x = rock_y, y = rock_r), color = "red", shape = 1) +  
  labs(x = "Year", y = "Rating", title = "Music Ratings by Genre") +  
  theme_minimal()
```

