



Majkowska Agata

Migdał-Najman Kamila

Najman Krzysztof

Raca Katarzyna

Analiza zgodności przekazu elementów werbalnych i niewerbalnych w wiadomościach tekstowych na Twitterze

www.ug.edu.pl



Problem i cel badawczy

Z badań językowych wynika, że przedstawiciele różnych grup wiekowych posługują się **odmiennym słownictwem** i formami gramatycznymi. Wydaje się, że różnicują je także powszechnie stosowane w wypowiedziach **znaki graficzne**, takie jak emotikony, emoji, piktogramy i inne znaki graficzne.

O ile zapis słowny wypowiedzi może być traktowany jako werbalna część wypowiedzi, to znaki graficzne będą ich niewerbalną częścią. Obie części wypowiedzi mogą być wzajemnie w różnych relacjach.

Celem prezentowanych badań jest próba oceny, który typ przekazu, werbalny czy niewerbalny, częściej występuje w wiadomościach na Twitterze i czy oba te typy przekazu są zgodne.

Autorzy podjęli próbę odpowiedzi na pytania: Jakie znaczenie uzyskują emoji w wiadomościach wysyłanych przez różne grupy wiekowe użytkowników? Czy identyfikacja najpopularniejszych emoji w wiadomości pozwala w przybliżeniu określić grupę wiekową autora wpisu?

Znaki graficzne - historia

```

19-Sep-82 11:44 Scott E Fahlman :-)
From: Scott E Fahlman <Fahlman at Cmu-20c>

I propose that the following character sequence for joke markers:

:-)

Read it sideways. Actually, it is probably more economical to mark
things that are NOT jokes, given current trends. For this, use

:-(
    
```

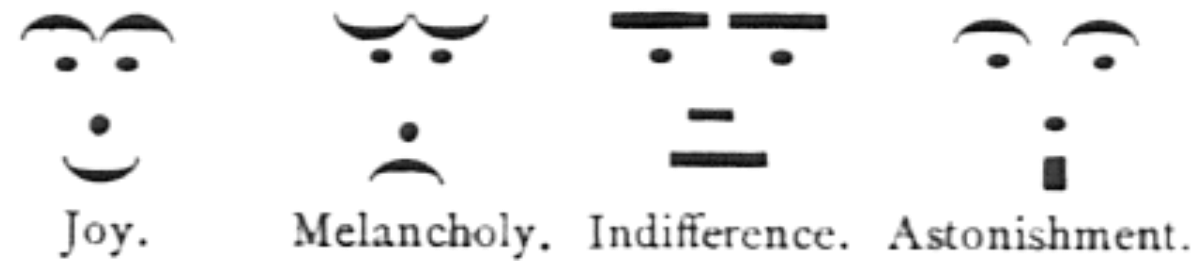


176 emoji



> 3 000 emoji

1881

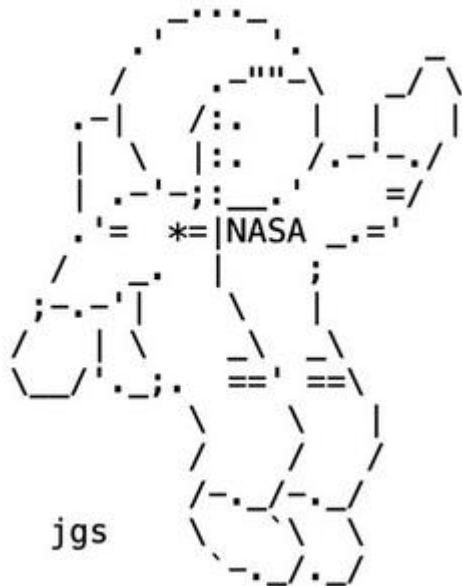


1982

KAOMOJI
 顔 “Kao” – twarz
 文字 “Moji” - postać

- | | |
|-----------|----------|
| Oczy: | Usta: |
| • * | • _ |
| • ^ | • . |
| • - | • O |
| <(^.^)> | <(^_ ^< |

1990

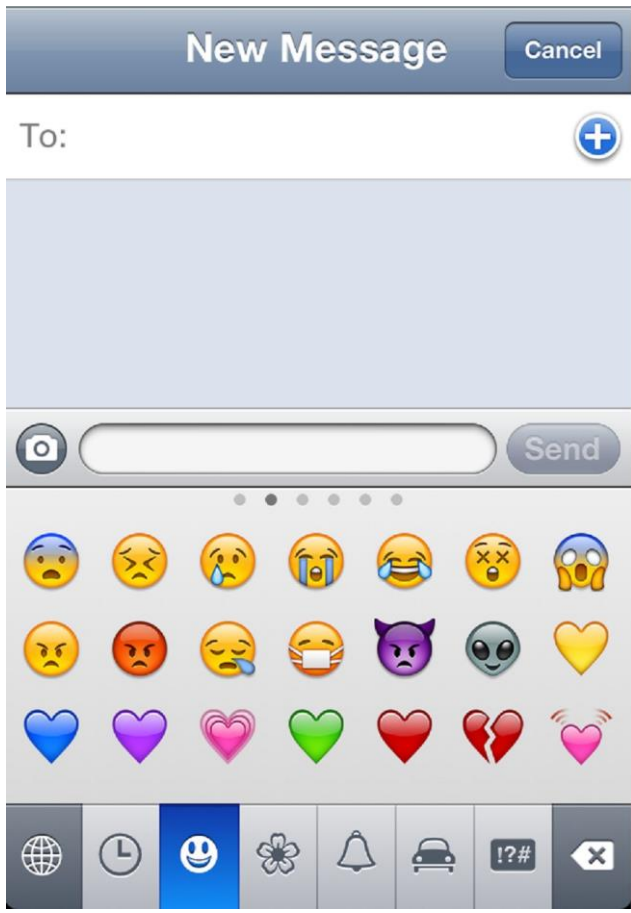


1999

2011



472 emoji



2022

PRZYGOTOWANIE DANYCH

DANE

100 NAJCZĘŚCIEJ WYSTĘPUJĄCYCH

SŁÓW W KAŻDEJ GRUPIE WIEKOWEJ
EMOJI W KAŻDEJ GRUPIE WIEKOWEJ

ŁĄCZNIE 535 UNIKALNYCH SŁÓW
652 UNIKALNYCH EMOJI

MACIERZ DANYCH

SŁÓW

535 ZMIENNYCH

UDZIAŁ WYSTĘPOWANIA DANEGO SŁOWA W 535
BADANYCH SŁOWACH ZE WZGLĘDU NA GRUPĘ WIEKOWĄ

DANE UNORMOWANE SUMUJĄ SIĘ DO 100% DLA GRUP WIEKOWYCH

67 PRZYPADKÓW

GRUPY WIEKOWE OD 13 DO 79 LATKÓW

EMOJI

652 ZMIENNYCH

UDZIAŁ WYSTĘPOWANIA DANEGO EMOJI W 652
BADANYCH EMOJI ZE WZGLĘDU NA GRUPĘ WIEKOWĄ

DANE UNORMOWANE SUMUJĄ SIĘ DO 100% DLA GRUP WIEKOWYCH

67 PRZYPADKÓW

GRUPY WIEKOWE OD 13 DO 79 LATKÓW

535 słów

652 emoji

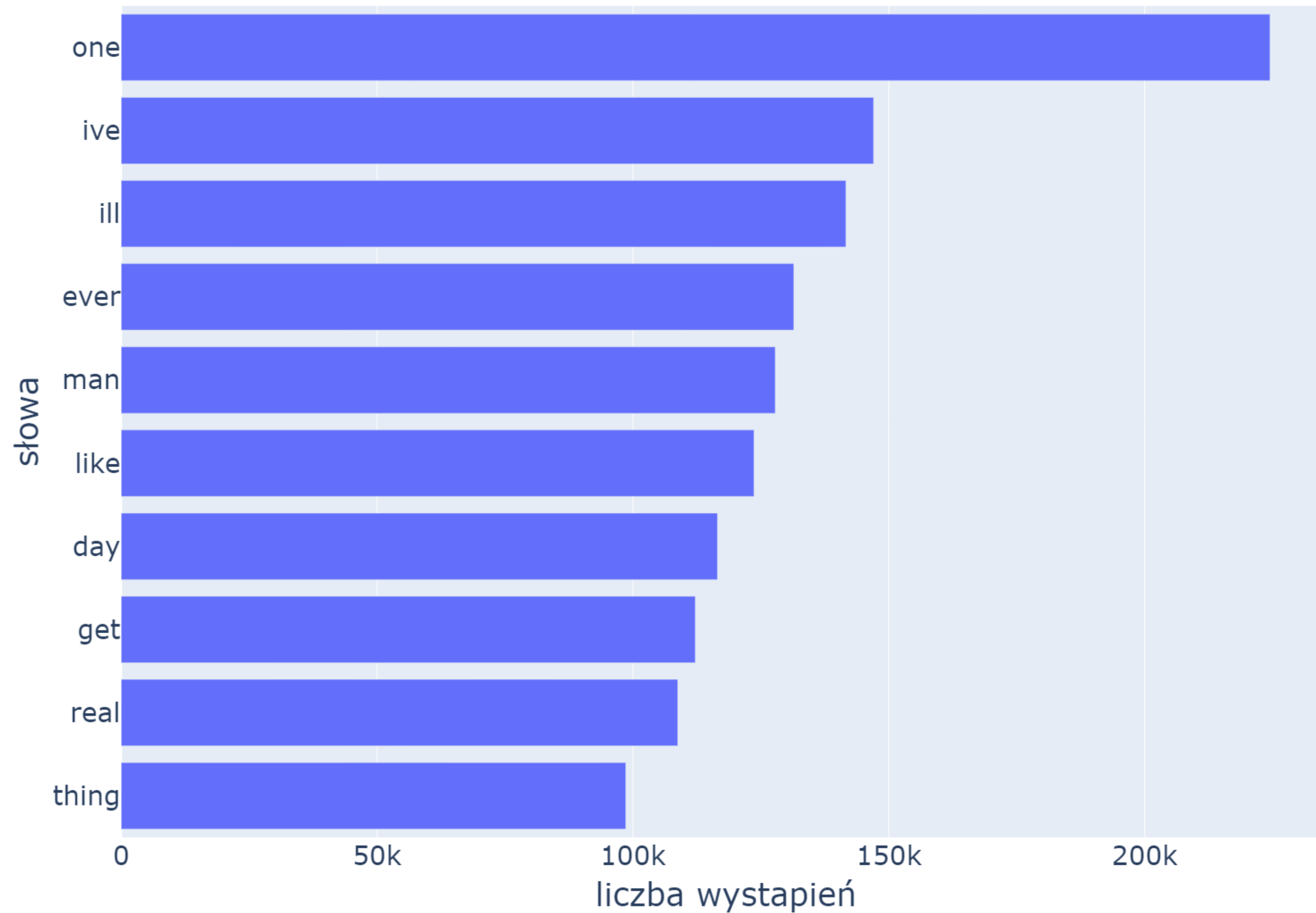
67 grup wiekowych

wiek	abigail	abushlomo	active	actually	agentneppi	airbrne	aku	album	alonma	also	
1	13	3	0	77	300	0	0	339	100	0	524
2	14	0	0	43	265	0	0	458	74	0	404
3	15	2	0	70	259	0	0	555	126	0	586
4	16	2	0	80	312	0	0	1007	97	0	626
5	17	3	0	90	396	0	1	593	122	0	586
6	18	6	0	107	405	0	0	945	218	0	726
7	19	3	0	169	566	0	0	1676	234	0	777
8	20	11	0	213	1267	0	0	811	599	0	1691
9	21	16	0	110	615	0	0	536	265	0	955
10	22	3	0	101	642	0	0	492	191	0	1005
11	23	0	0	112	683	0	0	511	263	0	1072
12	24	3	0	73	601	0	0	158	214	0	1056
13	25	2	0	82	605	0	0	208	276	0	906
14	26	5	0	68	495	0	0	226	150	0	850
15	27	2	0	58	472	0	0	220	175	1	760
16	28	2	0	63	307	0	0	122	89	0	664
17	29	1	0	49	338	0	0	220	63	0	519
18	30	0	0	74	569	0	0	131	177	0	1045
19	31	4	0	36	272	0	0	64	52	0	481
20	32	7	0	52	264	0	0	32	68	0	470

67 grup wiekowych

wiek	🍌	🌸	👩	🍷	🍷	🍌	👍	🌸	👩	🍌	...	👍	🍌	😬	👩	👍	👍	👍	👍	👍	Σ
13	35	0	3	2	2	1	12	12	0	32	...	1	1	24	7	0	0	0	1	20	16116
14	0	1	0	1	0	0	27	0	0	21	...	0	1	25	1	0	0	0	2	50	17785
15	0	1	3	2	3	0	7	8	0	38	...	12	0	13	2	0	0	0	0	112	22802
16	2	0	1	0	5	0	28	7	0	16	...	0	0	19	3	4	55	0	2	69	24822
17	8	0	8	2	3	1	7	30	0	101	...	2	1	29	1	0	2	0	2	133	28971
18	3	3	20	3	7	0	37	12	0	38	...	1	1	57	3	0	1	0	0	89	30053
19	10	5	2	2	7	1	127	46	0	42	...	0	0	28	3	0	0	1	10	117	37677
20	6	1	11	4	5	0	78	10	3	54	...	5	1	98	1	0	1	2	4	327	59314
21	52	0	43	10	4	0	31	11	0	33	...	1	1	25	4	3	2	0	9	223	37855
22	4	2	18	9	7	0	11	16	0	34	...	1	3	37	6	0	2	1	2	126	33659
23	11	3	11	3	9	0	24	3	0	128	...	1	0	69	3	0	1	1	2	138	35620
24	4	5	5	8	2	0	10	4	0	48	...	1	2	38	4	0	9	0	16	138	32877
25	60	1	29	7	6	1	55	7	0	53	...	11	15	62	15	1	1	0	5	110	34070

Najczęściej występujące słowa i emoji w tweetach



Emoji	Liczba wystąpień
😊	99 219
😭	63 815
❤️	42 316
☐	37 230
☐	29 421
😊	17 798
💓	16 006
☐	12 946
🔥	10 454
🙇	10 216

PRZYGOTOWANIE DANYCH

DANE

100 NAJCZĘŚCIEJ WYSTĘPUJĄCYCH

SŁÓW W KAŻDEJ GRUPIE WIEKOWEJ
EMOJI W KAŻDEJ GRUPIE WIEKOWEJ

ŁĄCZNIE 1187 ZMIENNYCH

535 UNIKALNYCH SŁÓW
652 UNIKALNYCH EMOJI

MACIERZ DANYCH

1187 ZMIENNYCH

UDZIAŁ WYSTĘPOWANIA DANEGO SŁOWA, EMOJI W 1187
BADANYCH SŁÓWACH I EMOJI ZE WZGLĘDU NA GRUPĘ
WIEKOWĄ

DANE UNORMOWANE SUMUJĄ SIĘ DO 100% DLA GRUP WIEKOWYCH

67 PRZYPADKÓW

GRUPY WIEKOWE OD 13 DO 79 LATKÓW

652 emoji + 535 słów

652 emoji + 535 słów

UNORMOWANIE

67 grup wiekowych

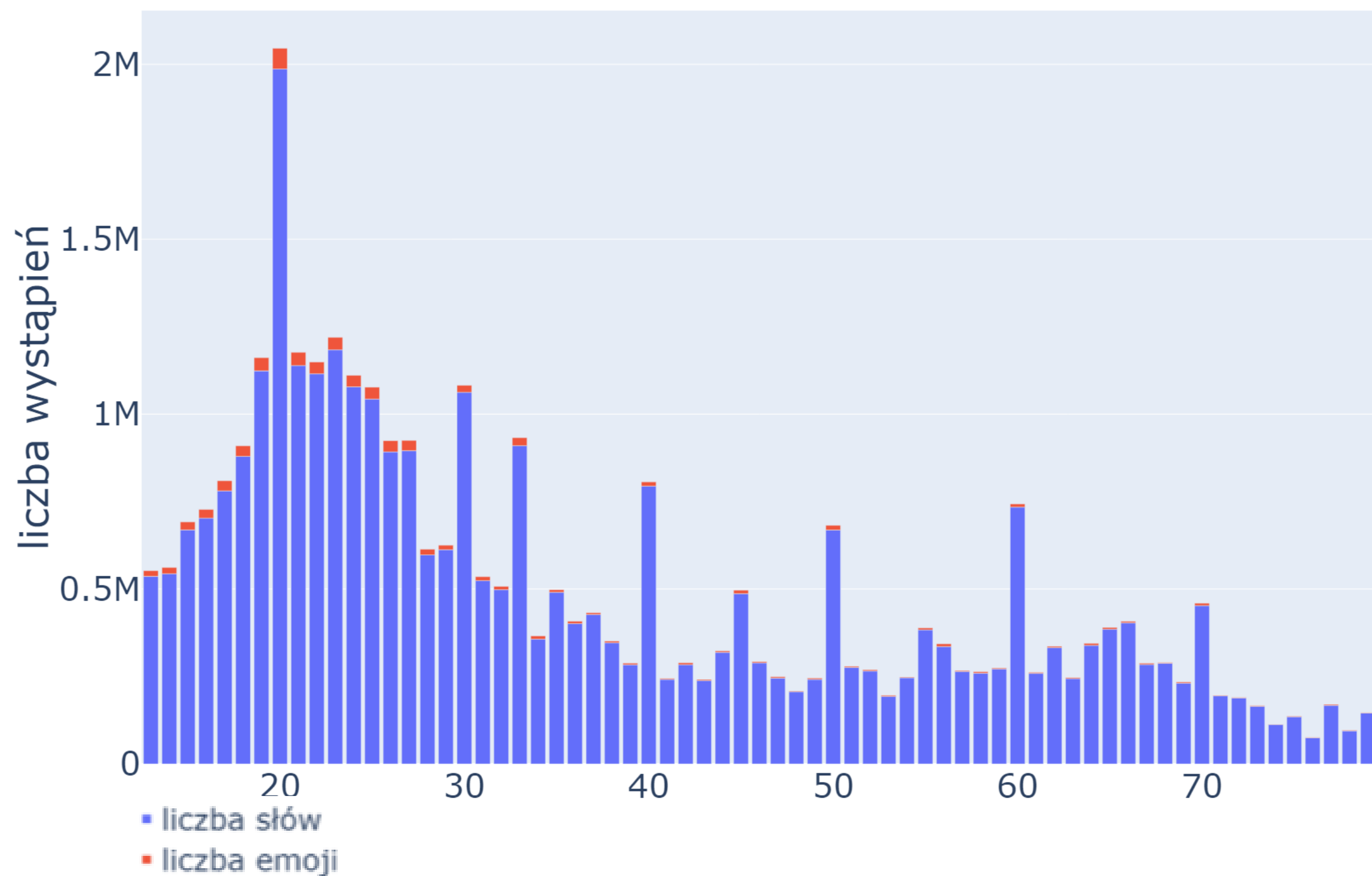
wiek	🍅	🌻	👶	🍷	🍷	🍷	👍	🌸	👶	👉	...	👉	🍷	👶	👉	👉	👉	👉	👉	Σ	
13	35	0	3	2	2	1	12	12	0	32	...	1	1	24	7	0	0	0	1	20	16116
14	0	1	0	1	0	0	27	0	0	21	...	0	1	25	1	0	0	0	2	50	17785
15	0	1	3	2	3	0	7	8	0	38	...	12	0	13	2	0	0	0	0	112	22802
16	2	0	1	0	5	0	28	7	0	16	...	0	0	19	3	4	55	0	2	69	24822
17	8	0	8	2	3	1	7	30	0	101	...	2	1	29	1	0	2	0	2	133	28971
18	3	3	20	3	7	0	37	12	0	38	...	1	1	57	3	0	1	0	0	89	30053
19	10	5	2	2	7	1	127	46	0	42	...	0	0	28	3	0	0	1	10	117	37677
20	6	1	11	4	5	0	78	10	3	54	...	5	1	98	1	0	1	2	4	327	59314
21	52	0	43	10	4	0	31	11	0	33	...	1	1	25	4	3	2	0	9	223	37855
22	4	2	18	9	7	0	11	16	0	34	...	1	3	37	6	0	2	1	2	126	33659
23	11	3	11	3	9	0	24	3	0	128	...	1	0	69	3	0	1	1	2	138	35620
24	4	5	5	8	2	0	10	4	0	48	...	1	2	38	4	0	9	0	16	138	32877
25	60	1	29	7	6	1	55	7	0	53	...	11	15	62	15	1	1	0	5	110	34070

67 grup wiekowych

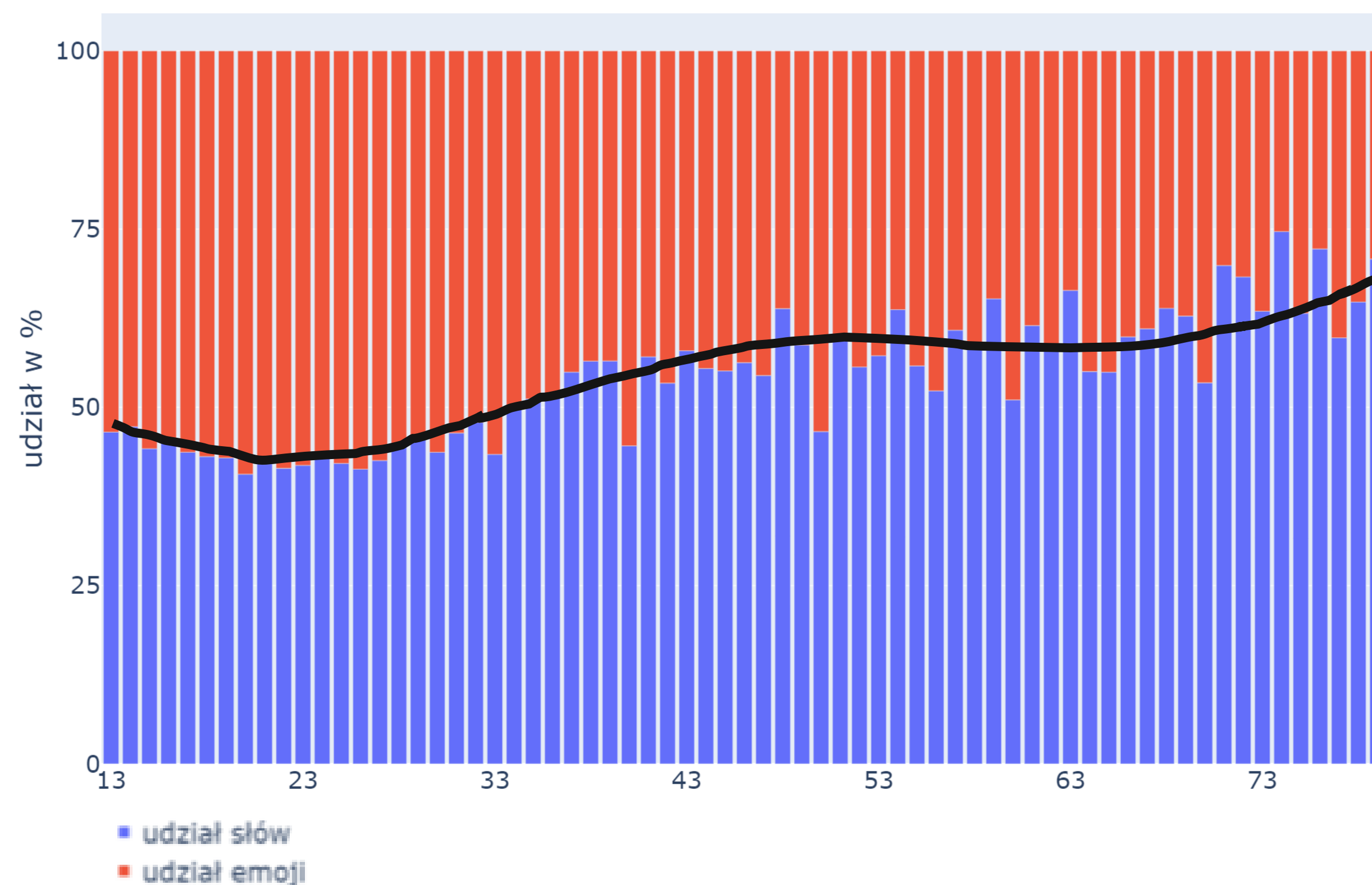
wiek	🍅	🌻	👶	🍷	🍷	🍷	👍	🌸	👶	👉	...	👉	🍷	👶	👉	👉	👉	👉	👉	Σ	
13	0.0022	0.0000	0.0002	0.0001	0.0001	0.0001	0.0007	0.0007	0.0000	0.0020	...	0.0001	0.0001	0.0007	0.0007	0.0000	0.0020	...	0.0001	0.0001	1.0
14	0.0000	0.0001	0.0000	0.0001	0.0000	0.0000	0.0015	0.0000	0.0000	0.0012	...	0.0000	0.0001	0.0000	0.0000	0.0012	...	0.0000	0.0001	0.0001	1.0
15	0.0000	0.0000	0.0001	0.0001	0.0001	0.0000	0.0003	0.0004	0.0000	0.0017	...	0.0005	0.0001	0.0000	0.0000	0.0017	...	0.0005	0.0001	0.0001	1.0
16	0.0001	0.0000	0.0000	0.0000	0.0002	0.0000	0.0011	0.0003	0.0000	0.0006	...	0.0000	0.0001	0.0000	0.0006	...	0.0000	0.0001	0.0001	0.0001	1.0
17	0.0003	0.0000	0.0003	0.0001	0.0001	0.0000	0.0002	0.0010	0.0000	0.0035	...	0.0001	0.0001	0.0000	0.0035	...	0.0001	0.0001	0.0001	0.0001	1.0
18	0.0001	0.0001	0.0007	0.0001	0.0002	0.0000	0.0012	0.0004	0.0000	0.0013	...	0.0000	0.0001	0.0000	0.0013	...	0.0000	0.0001	0.0001	0.0001	1.0
19	0.0003	0.0001	0.0001	0.0001	0.0002	0.0000	0.0034	0.0012	0.0000	0.0011	...	0.0000	0.0001	0.0000	0.0011	...	0.0000	0.0001	0.0001	0.0001	1.0
20	0.0001	0.0000	0.0002	0.0001	0.0001	0.0000	0.0013	0.0002	0.0001	0.0009	...	0.0001	0.0001	0.0000	0.0009	...	0.0001	0.0001	0.0001	0.0001	1.0
21	0.0014	0.0000	0.0011	0.0003	0.0001	0.0000	0.0008	0.0003	0.0000	0.0009	...	0.0000	0.0001	0.0000	0.0009	...	0.0000	0.0001	0.0001	0.0001	1.0
22	0.0001	0.0001	0.0005	0.0003	0.0002	0.0000	0.0003	0.0005	0.0000	0.0010	...	0.0000	0.0001	0.0000	0.0010	...	0.0000	0.0001	0.0001	0.0001	1.0
23	0.0003	0.0001	0.0003	0.0001	0.0003	0.0000	0.0007	0.0001	0.0000	0.0036	...	0.0000	0.0001	0.0000	0.0036	...	0.0000	0.0001	0.0001	0.0001	1.0
24	0.0001	0.0002	0.0002	0.0002	0.0001	0.0000	0.0003	0.0001	0.0000	0.0015	...	0.0000	0.0001	0.0000	0.0015	...	0.0000	0.0001	0.0001	0.0001	1.0
25	0.0018	0.0000	0.0009	0.0002	0.0002	0.0000	0.0016	0.0002	0.0000	0.0016	...	0.0003	0.0001	0.0000	0.0016	...	0.0003	0.0001	0.0001	0.0001	1.0

Podstawowe statystyki używania słów i emoji

Liczba wystąpień słów i emoji w tweetach dla grup wiekowych



Udział wystąpień różnych słów i emoji w tweetach dla grup wiekowych



Najczęściej występujące słowa i emoji w analizowanych tweetach użytkowników Twittera.

one	ive	ill	ever	man	like	day	get	real	😄
224 480	147 005	141 610	131 422	127 799	123 647	116 504	112 167	108 748	99 219

ANALIZA SKUPIEŃ

METODA WIĄZANIA ZOSTAŁA WYBRANA
W OPARCIU O WSPÓŁCZYNNIK
KORELACJI KOFENETYCZNEJ

WYKORZYSTANA ZOSTAŁA HIERARCHICZNA METODA
GRUPOWANIA

WZÓR:

$$r_{kof} = \frac{\sum_{r,s;r < s} (d_{rs} - \bar{d})(c_{rs} - \bar{c})}{\sqrt{\sum_{r,s;r < s} (d_{rs} - \bar{d})^2 \cdot \sum_{r,s;r < s} (c_{rs} - \bar{c})^2}}$$

ODLEGŁOŚĆ

ZE WZGLĘDU NA WYSTĘPOWANIE PROFILI WZGLĘDNYCH
ZASTOSOWANO ODLEGŁOŚĆ OPARTĄ
NA PROPORCJACH (WSPÓŁCZYNNIK
LOKALIZACJI FLORENCE'A)

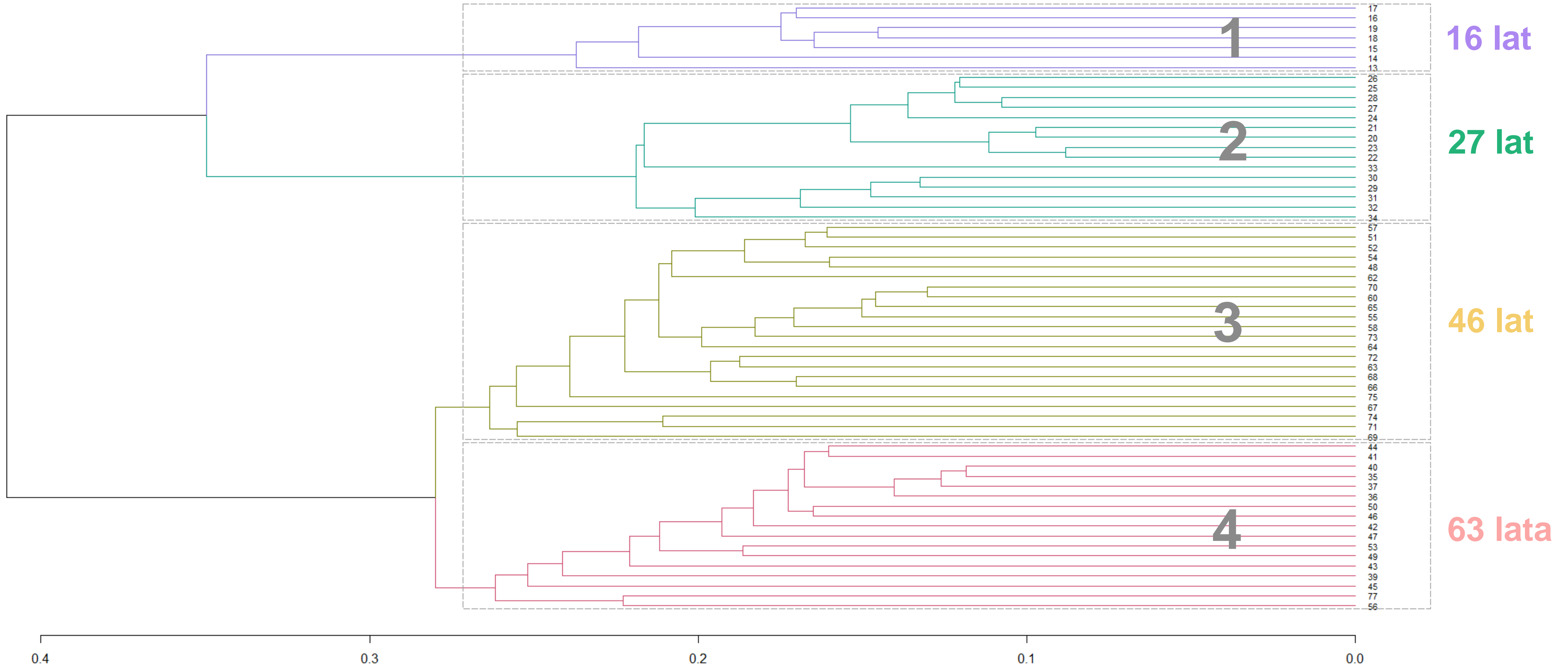
WZÓR:

$$d_{rs} = \frac{1}{200} \sum_{j=1}^p |p_{i,j} - p_{i,k}|$$

Metoda łączenia	Współczynnik korelacji kofenetycznej
najbliższego sąsiada	0,5776
najdalszego sąsiada	0,7274
średniej grupowej	0,7344
mediany	0,3750
centroidalna	0,6976

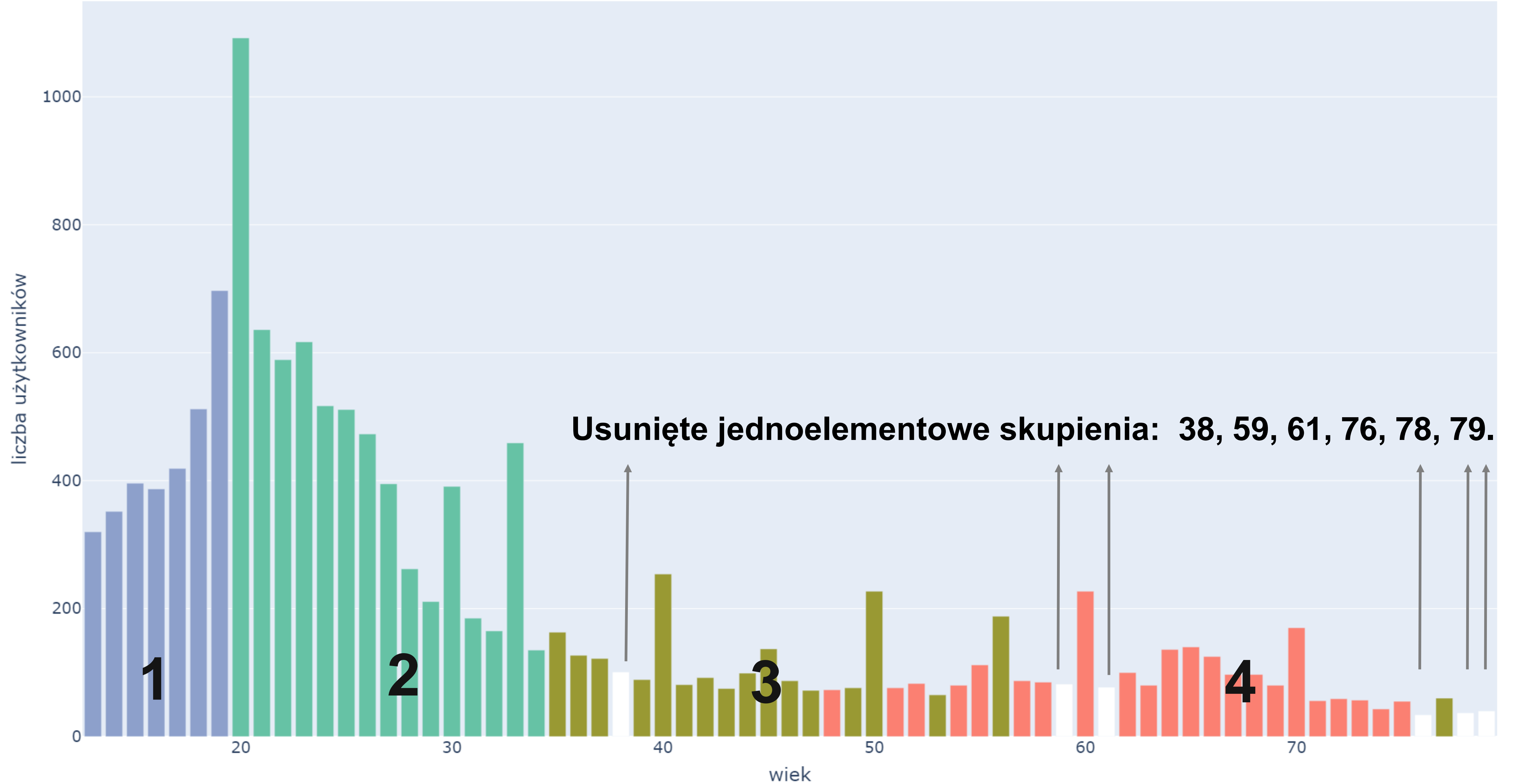
Struktura skupień badanych 67 grup wiekowych

Średni wiek
w skupieniu



Z dendrogramu zostały usunięte jednoelementowe skupienia: 38, 59, 61, 76, 78, 79 lat.

Struktura skupień badanych 67 grup wiekowych



Podstawowe charakterystyki wyodrębnionych skupień

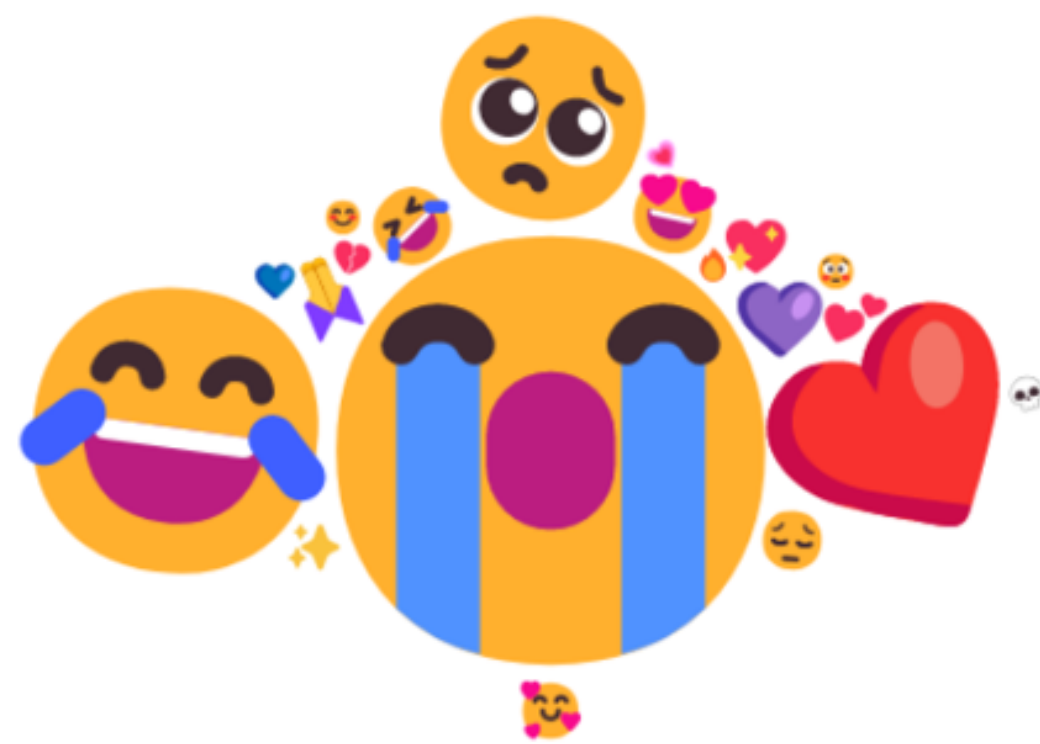
Skupienie	Liczba grup wiekowych	Liczba użytych różnych słów i emoji z 1 187 wyodrębnionych do analizy	Udział użytych słów i emoji z 1 187 wyodrębnionych do analizy
1	7	992	83,57%
2	15	1 040	87,62%
3	17	1 052	88,63%
4	22	1 022	86,10%
5	1	678	57,12%
6	1	581	48,95%
7	1	623	52,49%
8	1	479	40,35%
9	1	511	43,05%
10	1	476	40,10%

skupienia jednoelementowe

SKUPIENIE 1

KOMUNIKACJA WERBALNA

KOMUNIKACJA NIEWERBALNA



POWIĄZANIA MIĘDZY KOMUNIKACJĄ WERBALNĄ I NIEWERBALNĄ

SŁOWO → EMOJI

EMOJI → SŁOWO

One → 😭

😭 → Like

Like → 😭

😭 → Lol

ive → 😊

❤️ → Love

Man → 😭

😞 → Thank

Ill → 😭

💜 → Love

Ever → 😭

🙏 → Please

Day → 😞

😍 → Love

Love → 😞

😜 → Funny

Get → 😭

❤️ → Love

Dont → 😭

😓 → Sorry

SKUPIENIE 2

KOMUNIKACJA WERBALNA

KOMUNIKACJA NIEWERBALNA



POWIĄZANIA MIĘDZY KOMUNIKACJĄ WERBALNĄ I NIEWERBALNĄ

SŁOWO → EMOJI

EMOJI → SŁOWO

One → 😊

😊 → Lol

Ive → 😭

😭 → Much

Like → 😊

❤️ → Love

Ever → 😊

😜 → Like

Ill → 😊

😞 → Thank

Day → 😊

😍 → Love

Get → 😊

😄 → Love

Man → 😊

💜 → Love

Thing → 😊

🔥 → Fire

Dont → 😊

❤️ → Love

SKUPIENIE 3

KOMUNIKACJA WERBALNA



KOMUNIKACJA NIEWERBALNA



POWIĄZANIA MIĘDZY KOMUNIKACJĄ WERBALNĄ I NIEWERBALNĄ

SŁOWO → EMOJI

One → 😂

Ill → 😂

Ive → 😂

Man → 😂

Ever → 😂

Day → 🍰

Real → 😂

Like → 😂

Get → 😂😂

Thing → 😂

EMOJI → SŁOWO

😂 → Lol

😂 → Lmao

❤️ → Love

😭 → Want, Much

💜 → Thank

😍 → Love

🔥 → Fire

🌹 → تحياتي ("moje pozdrowienia")

👍 → Mate

😊 → Thank

SKUPIENIE 4

KOMUNIKACJA WERBALNA



KOMUNIKACJA NIEWERBALNA



POWIĄZANIE MIĘDZY KOMUNIKACJĄ WERBALNĄ I NIEWERBALNĄ

SŁOWO → EMOJI

Trump → 😂

One → 😂

Real → 😂😂

Ill → 😂

Ive → 😂

Man → 😂

Ever → ❤️

Get → 😂

Donald → 🙏😭👍us🗨️🤪

Thing → 😂

EMOJI → SŁOWO

😂 → Lol

😂 → Good

❤️ → Love

us → Usa

😭 → Harassment, Thank

🙏 → Amen

👍 → Good

💜 → Love

😭 → Duh

🙏 → Thank

Wnioski końcowe

- Wyróżniono 4 główne skupienia grup wiekowych. Przeciętny wiek w wyróżnionych grupach jest następujący: 16 lat, 27 lat, 46 lat i 63 lata.
- Wyróżnione skupienia są względnie homogeniczne co do występowania zarówno słów i emoji w wiadomościach publikowanych na Twitterze.
- Największe podobieństwo stosowanych słów i emoji w wiadomościach na Twitterze zaobserwowano w dwóch najmłodszych grupach wiekowych: 13-19 lat oraz 20-34 lata.
- Grupy wiekowe między 48 a 56 rokiem życia przeplatają się pomiędzy sąsiadującymi skupieniami.
- Między skupieniami można zaobserwować istotne różnice w strukturze używanych słów i emoji.
- Młodzi częściej piszą o swoich własnych sprawach, miłości i zabawie, starsi o sprawach dotyczących społeczeństwa, o polityce, zdrowiu i bogu.

Wnioski końcowe

- Przekaz niewerbalny reprezentowany w badaniu przez emoji jest nie mniej, a może nawet bardziej ważny od werbalnego.
- Zróżnicowanie emoji jest większe niż słów (652 emoji, 535 słów). Jedno emoji potrafi przekazać więcej informacji niż jedno słowo choć w bardziej złożony sposób.
- Najciekawsze są wyniki analizy sposobu łączenia się słów z emoji i emoji ze słowami. Emoji dużo bardziej różnicowały słowa, z którymi razem współwystępowały, niż słowa, które współwystępowały z różnymi emoji.
- Umieszczenie emoji w punkcie centralnym, a więc słowa obok emoji, pozwoliło na uzyskanie większej emfazy i emocjonalności wiadomości tekstowej. A tym samym emoji w połączeniu ze słowem może stać się dominującym nośnikiem treści. Wydaje się więc, że to emoji, czyli przekaz niewerbalny pozwala w większym stopniu na identyfikację grupy wiekowej autora wpisu. Przekaz niewerbalny staje się przekazem samym w sobie, a słowa zaczynają pełnić rolę uzupełniającą treść przekazu.

Jednak w tym zakresie potrzebne są jeszcze kolejne badania.



Majkowska Agata

Migdał-Najman Kamila

Najman Krzysztof

Raca Katarzyna

Analiza zgodności przekazu elementów werbalnych i niewerbalnych w wiadomościach tekstowych na Twitterze

www.ug.edu.pl

