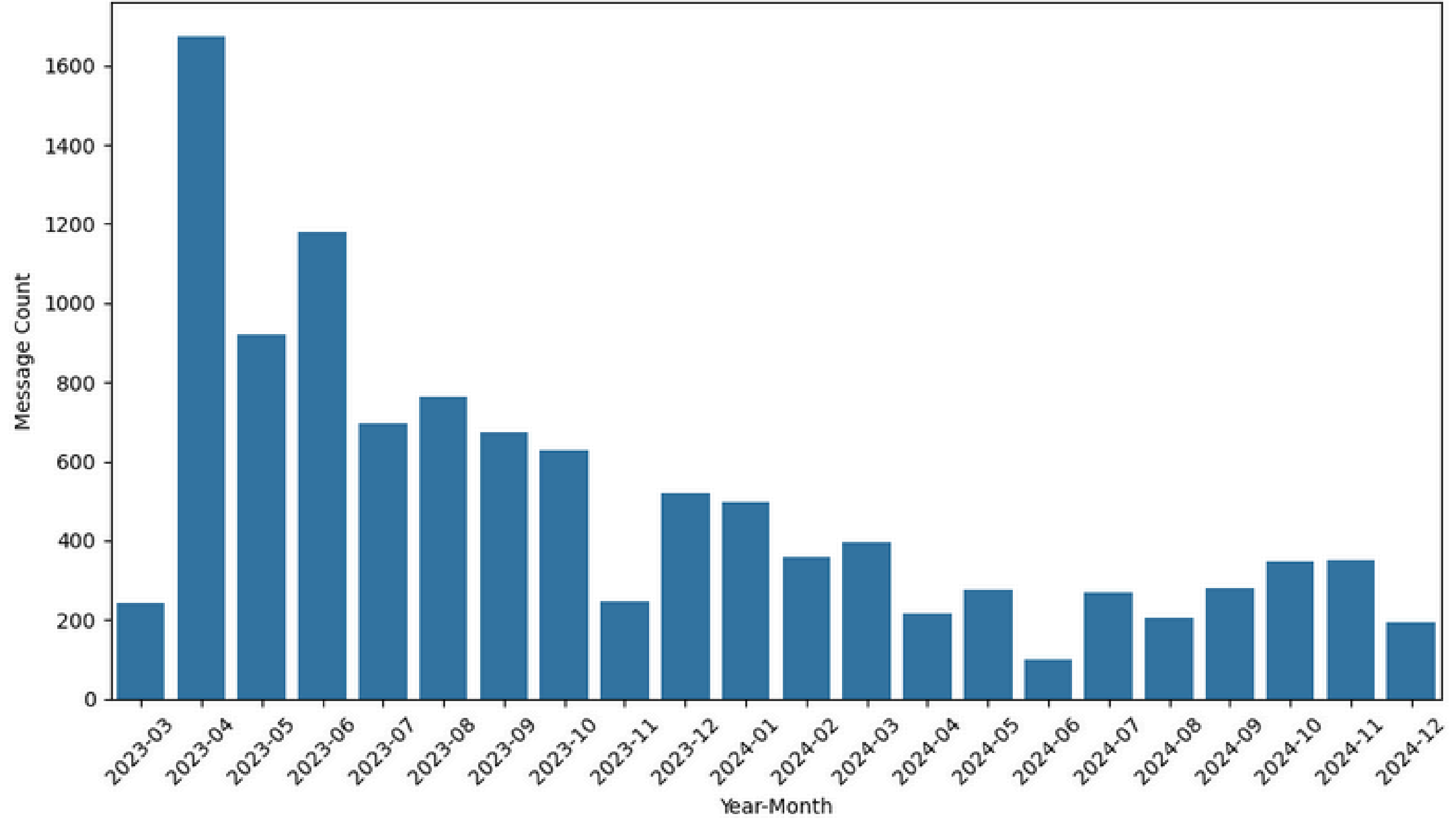


Message Count for a Contact



whatsapp **love** stories ♥

Juanjo **MONTERO**

mentor: Yunier **CORDOVA**

ITAcademy 2025

WhatsApp Chat with Alguien.txt X

```
1 10/29/21, 10:18 - 123456789: Te voy a buscar? Quedamos en Gracia?
2 10/29/21, 10:21 - Alguien: Pues como tú prefieras la verdad
3 10/29/21, 10:35 - 123456789: Luego vemos ;)
4 10/29/21, 12:42 - Alguien: <Media omitted>
5 10/29/21, 12:42 - Alguien: 🧑
6 10/29/21, 13:08 - 123456789: 😂😂😂😂
7 10/29/21, 13:08 - 123456789: Lo mejor del día
8 10/29/21, 13:47 - Alguien: 😂😂😂
9 10/29/21, 13:48 - Alguien: De momento... 😊
10 10/29/21, 17:44 - 123456789: Te paso a buscar por el curro y ya vemos como vamos.
```

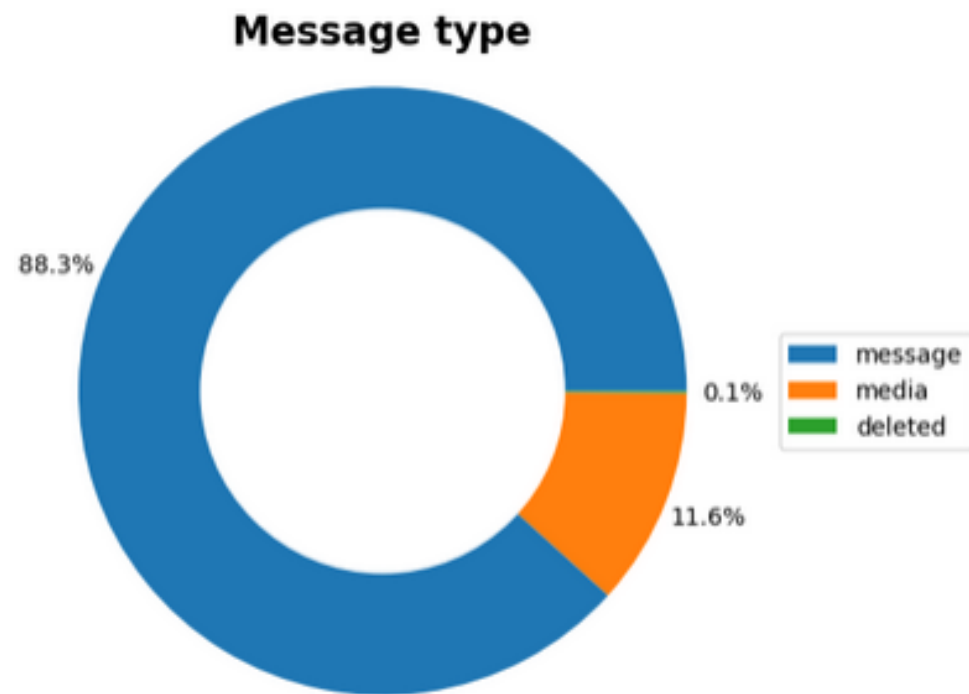
Exported WhatsApp backup files

52 different txt files - around 200k messages

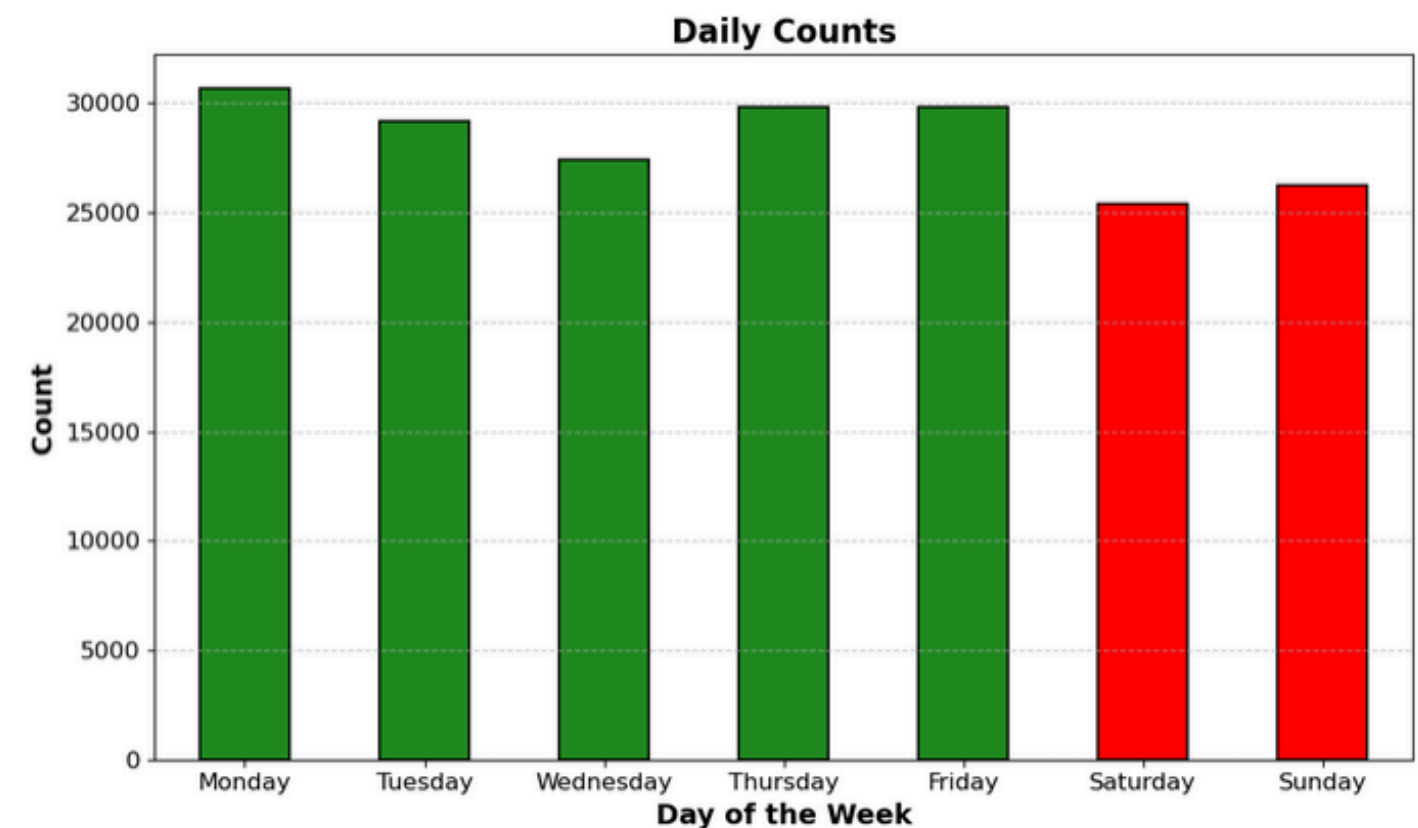
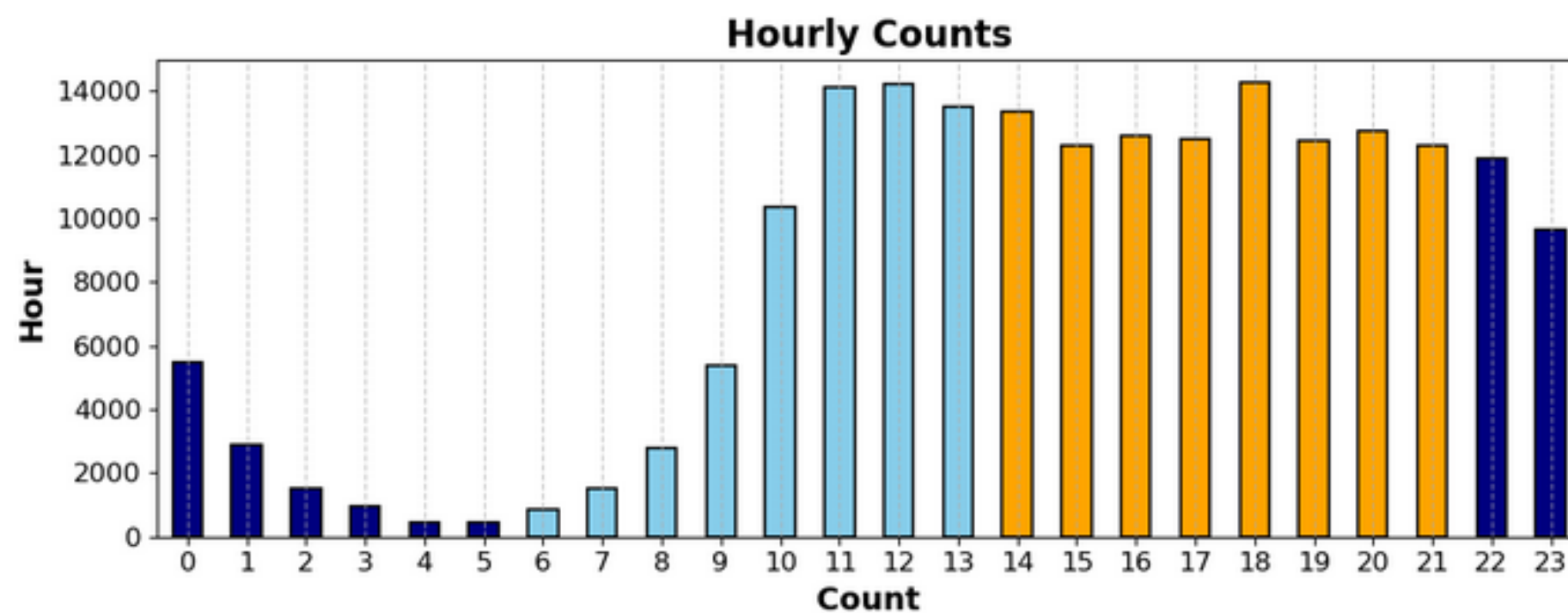


CERES	2022-03-04 14:34:00	ME	CERES	Nah, par de horas basta.
CERES	2022-03-04 14:34:00	CERES	ME	Poke te hace pues?
CERES	2022-03-04 14:34:00	ME	CERES	Sii 🤔🤔
CERES	2022-03-04 14:34:00	CERES	ME	No, lo compro
CERES	2022-03-04 14:36:00	CERES	ME	Pero el pescado lo compro congelado?
CERES	2022-03-04 14:37:00	ME	CERES	Es que debe haber estado congelado! Por el ani...
CERES	2022-03-04 14:39:00	CERES	ME	Ya, pero no sé si se descongelara s tiempo
CERES	2022-03-04 16:19:00	CERES	ME	Ya estoy

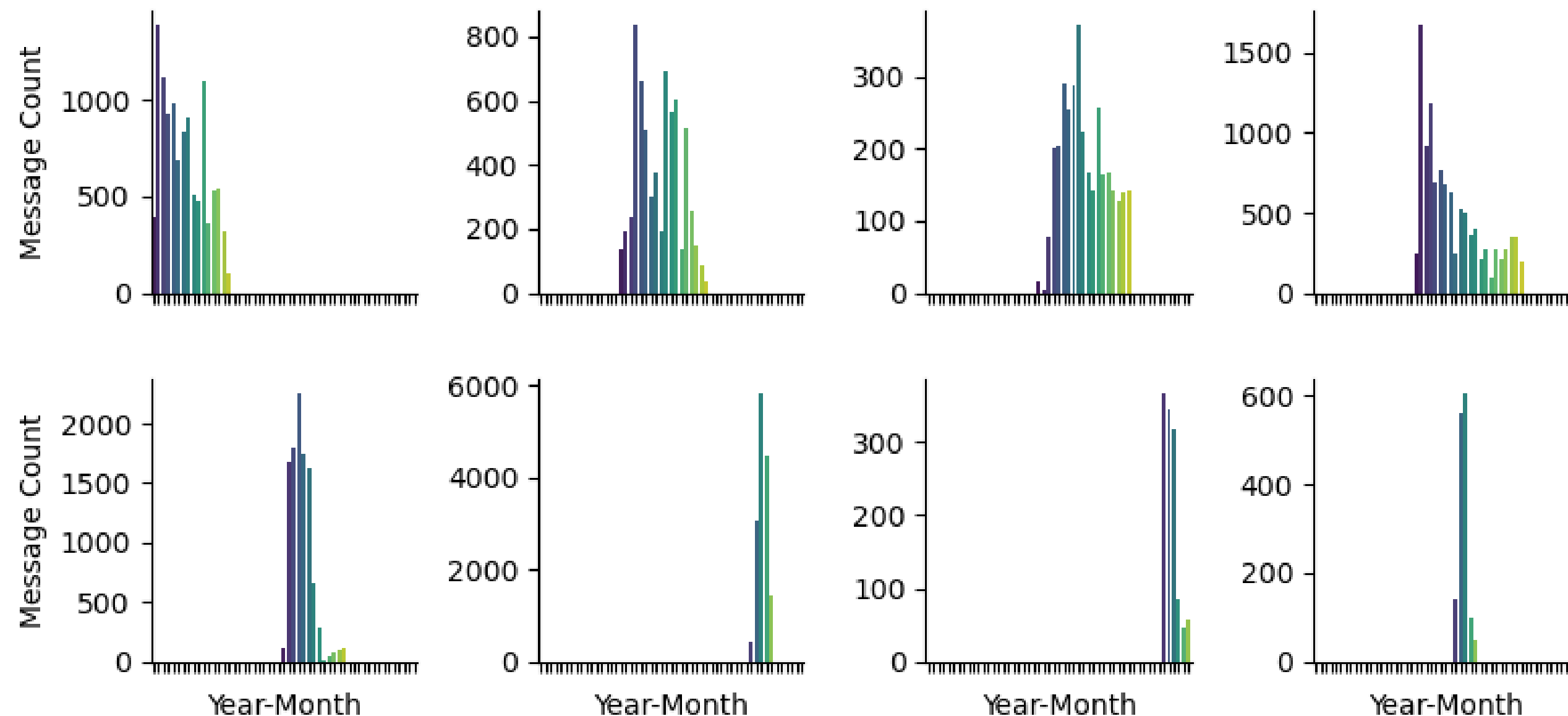
initial exploration



media can be voicenotes, pictures, videos and GIFs

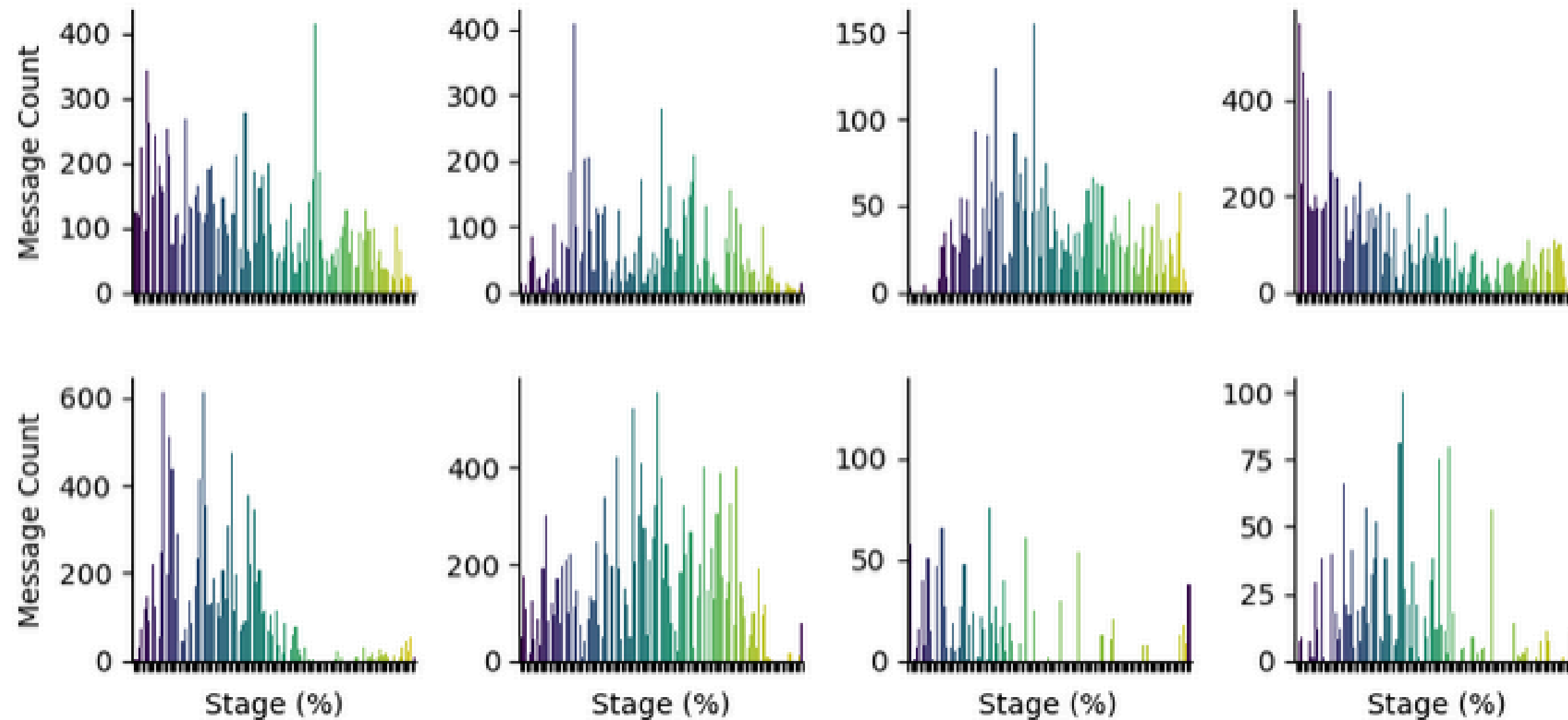


message count over time



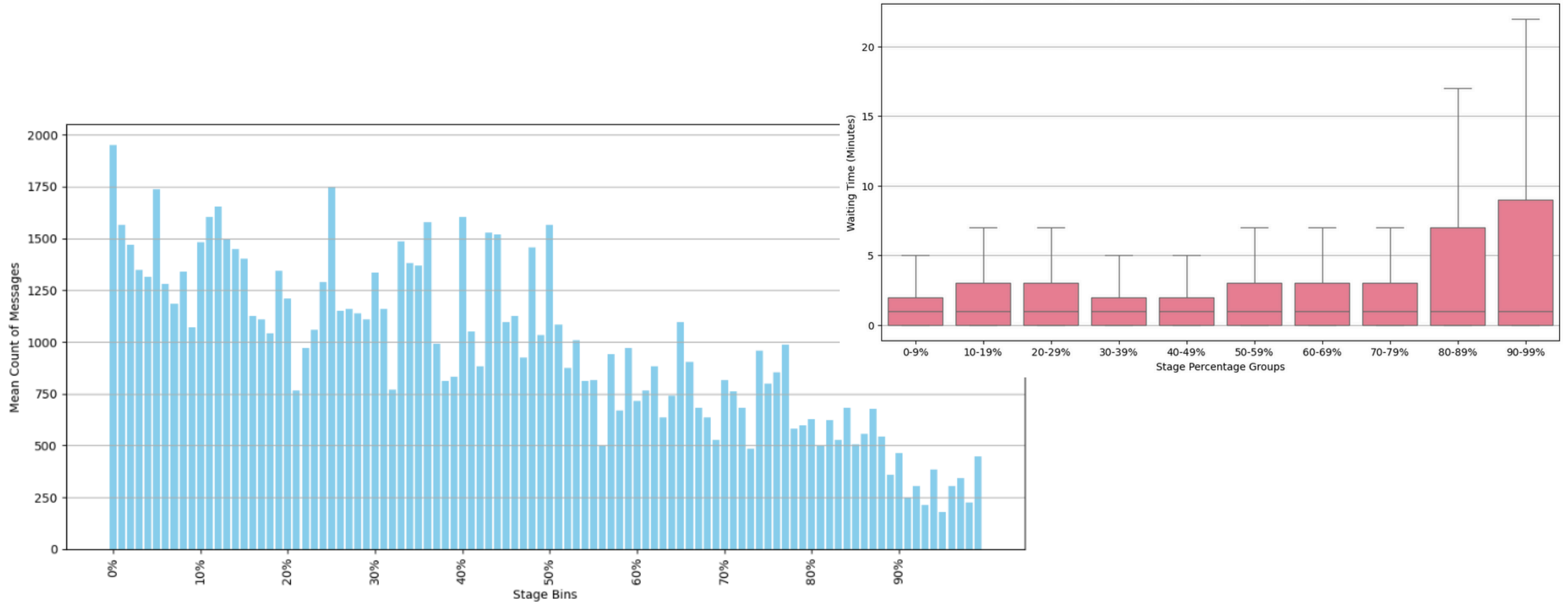
for 8 randomly selected contacts out of 16

message count from the beginning to the (bitter) end



for the same randomly selected contacts

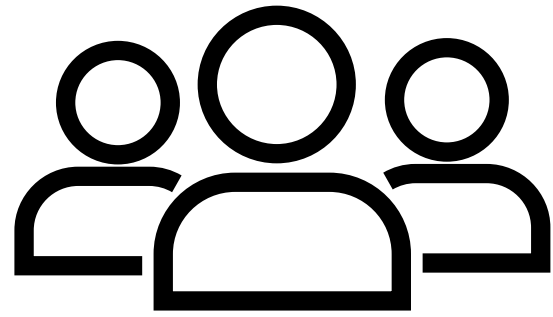
waiting time at different stages



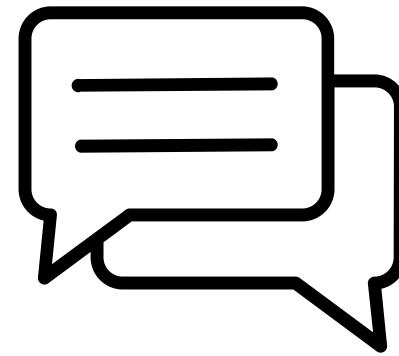
mean message count for all contacts

with the length of the relationship normalized

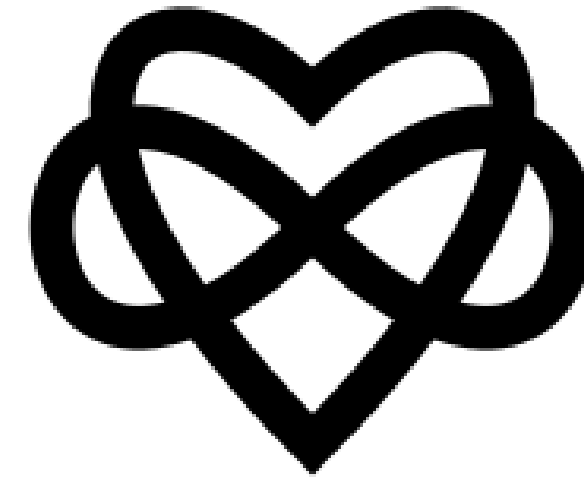
ethical non-monogamy



16 contacts

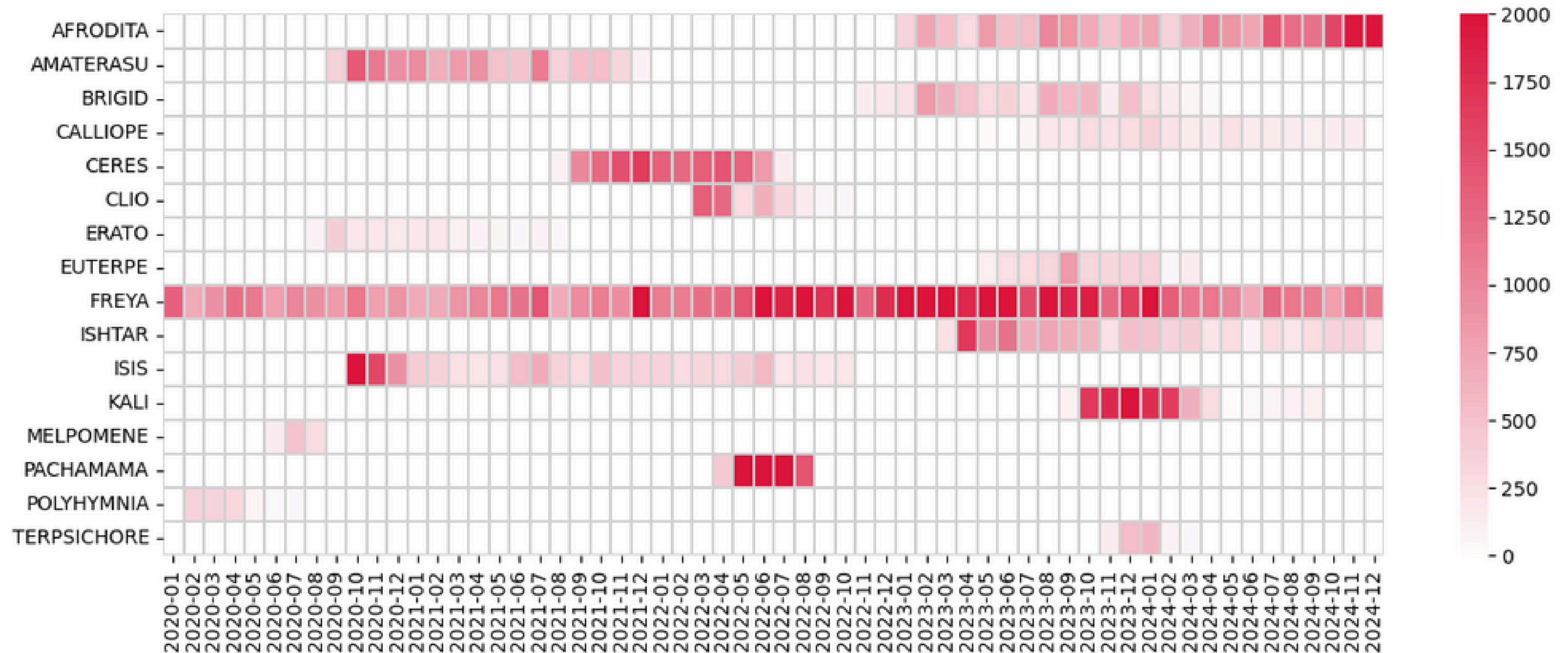


~200k



In contemporary discourse, non-monogamy is understood to encompass a variety of practices, ranging from culturally institutionalised arrangements like polygamy to consensual agreements such as **open relationships and polyamory**.

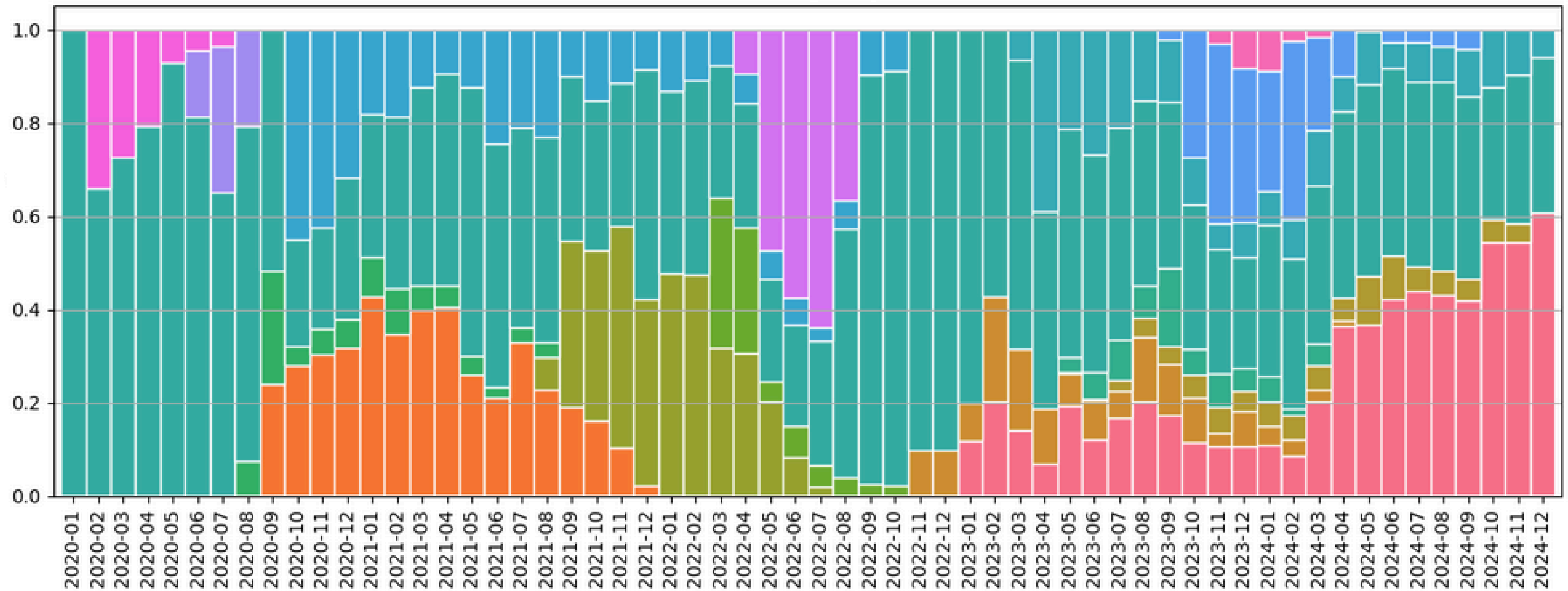
monthly **message count** for each contact



5 years period (2020-2024)

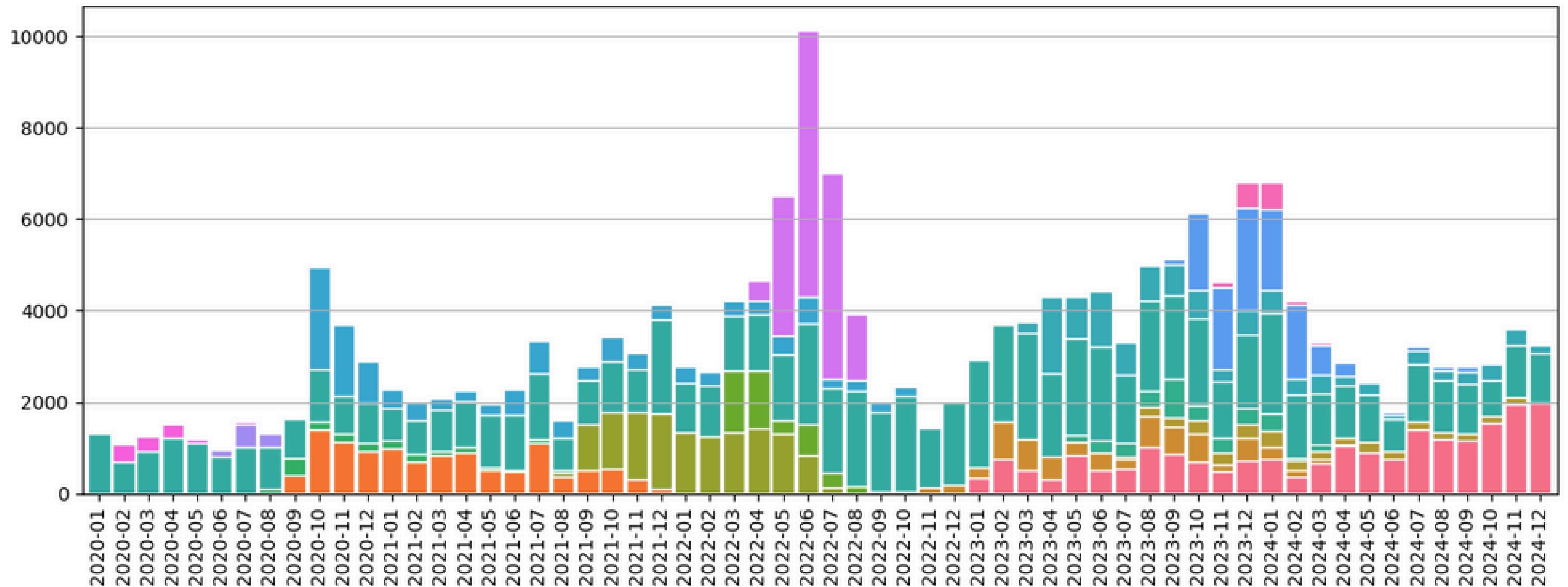
what about **sharing**?

message ratio by contact, month after month

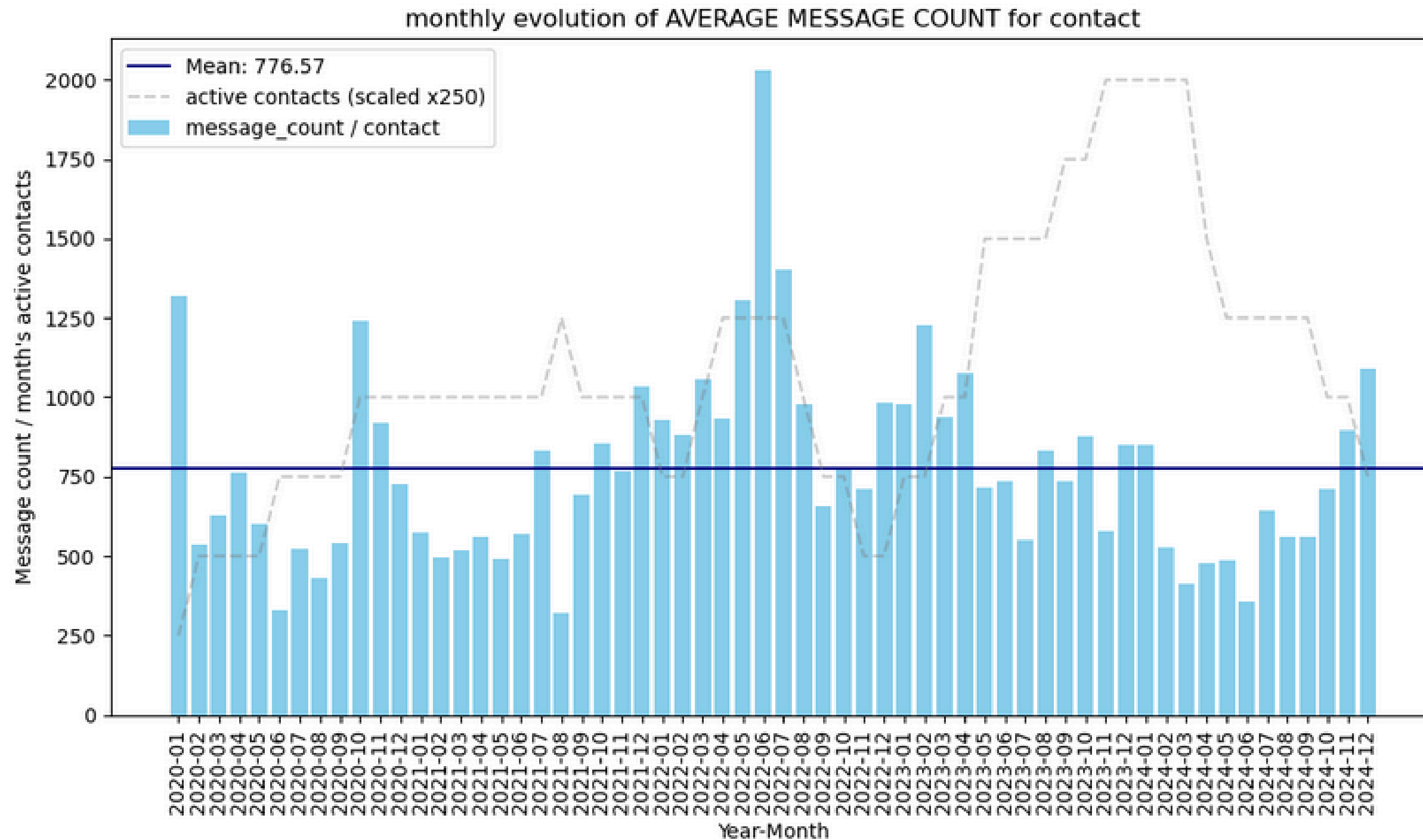


what about **sharing**?

stacked message count by contact, month after month



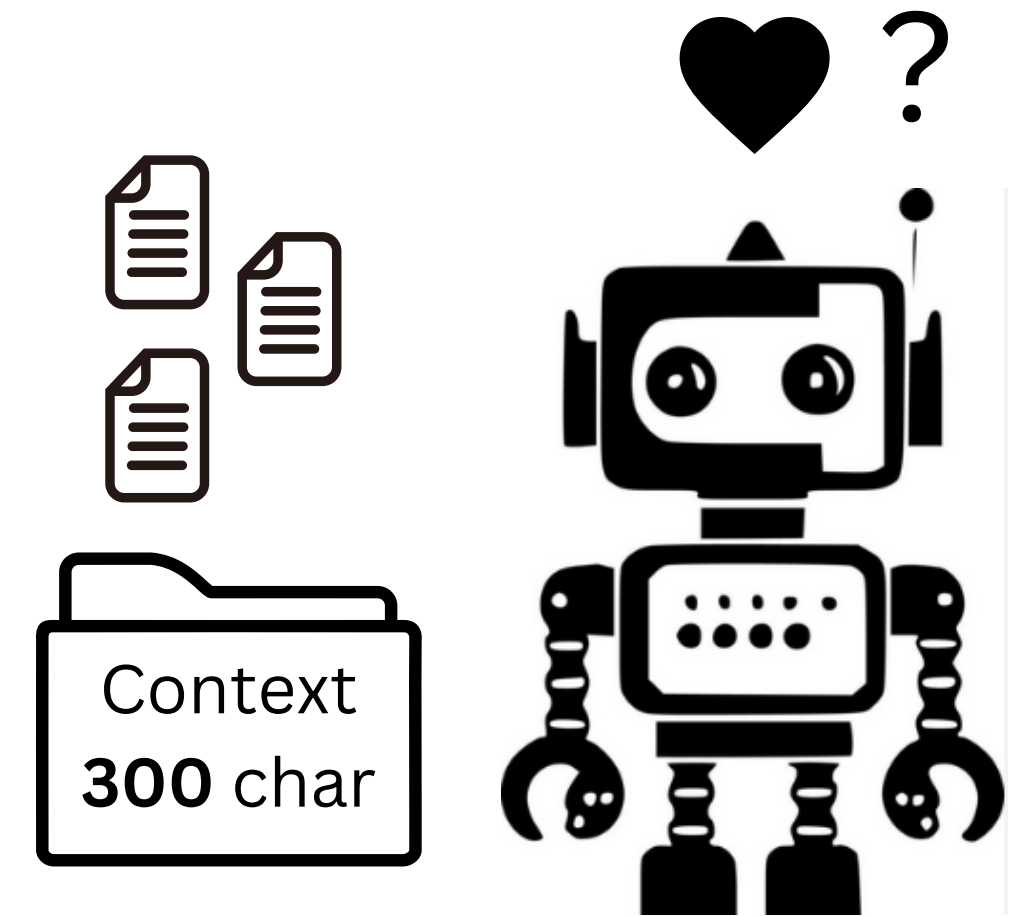
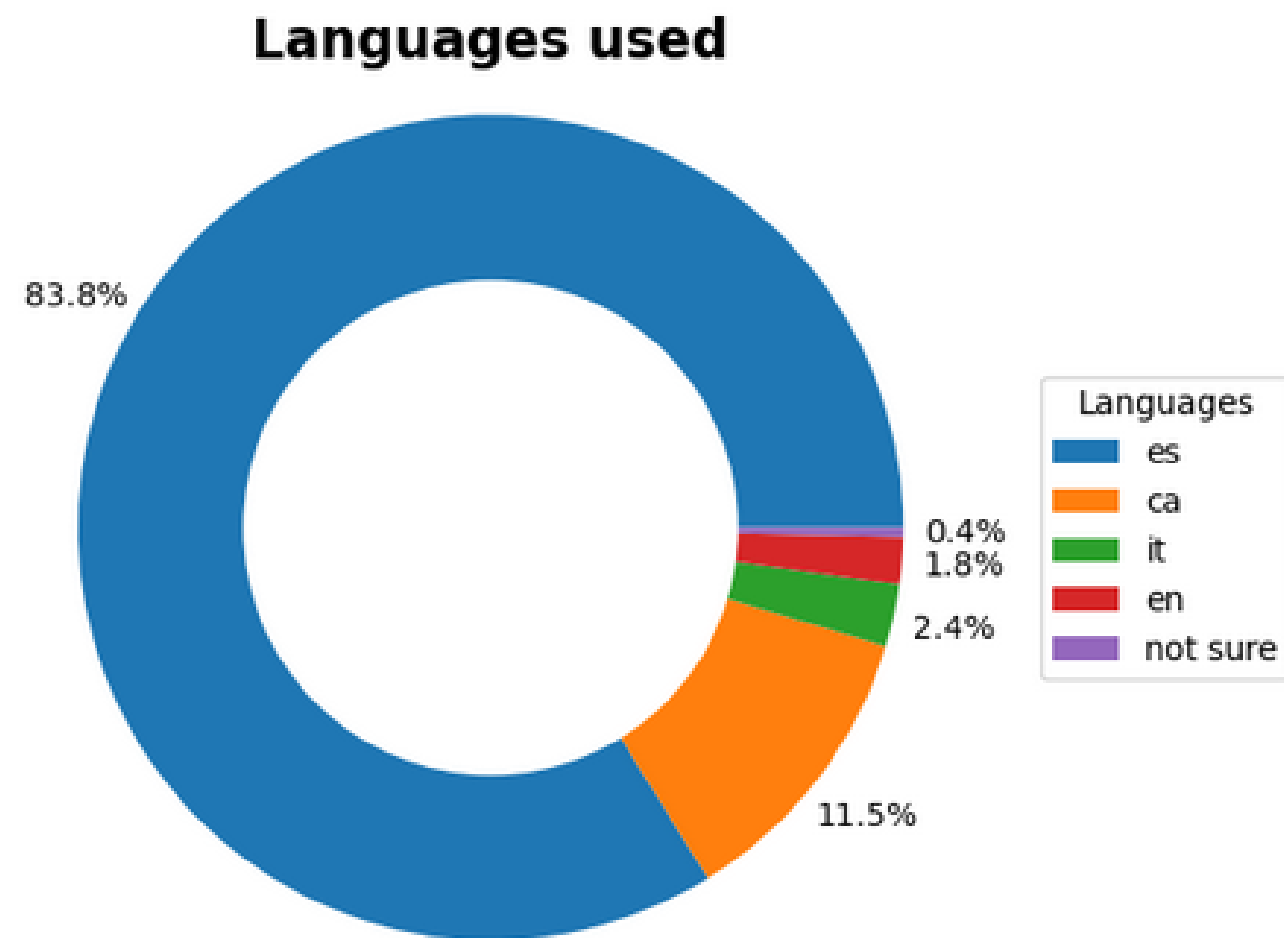
how many messages does **each contact** get?



Correlation between average monthly messages per contact and active monthly contacts: -0.11

the message should be **in the message**

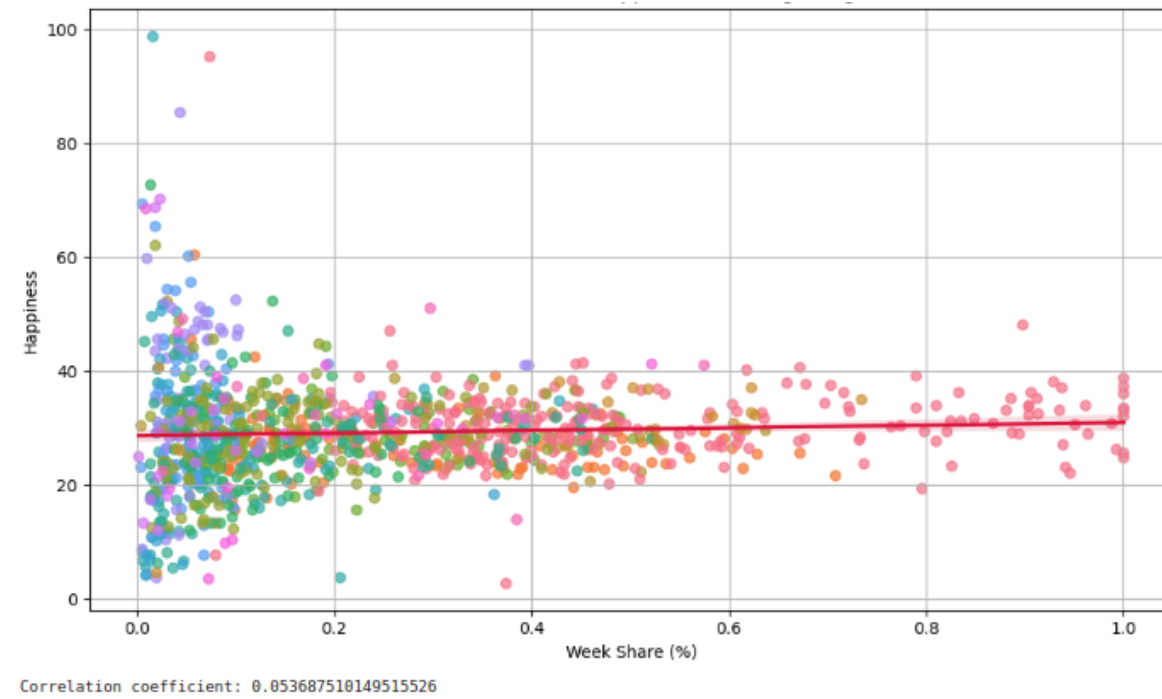
polyglot for **language detection**



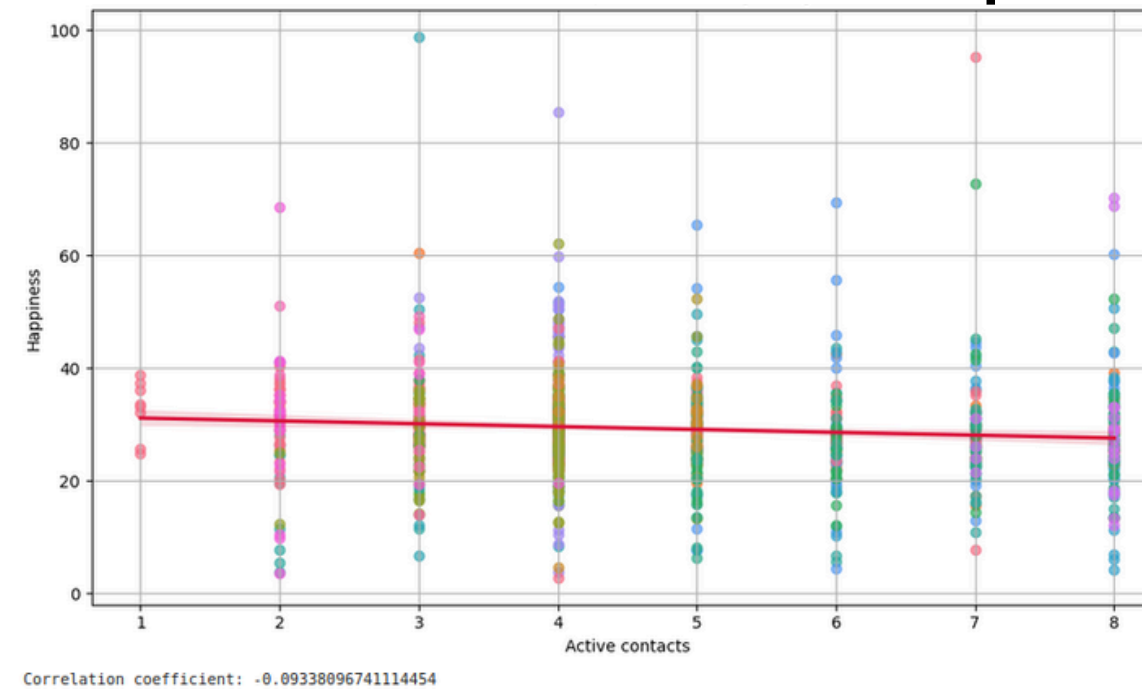
transformers **sentiment analysis** pipeline,
BERTopic for topic analysis
both over **multilingual BERT**

what happens with **happiness** as...

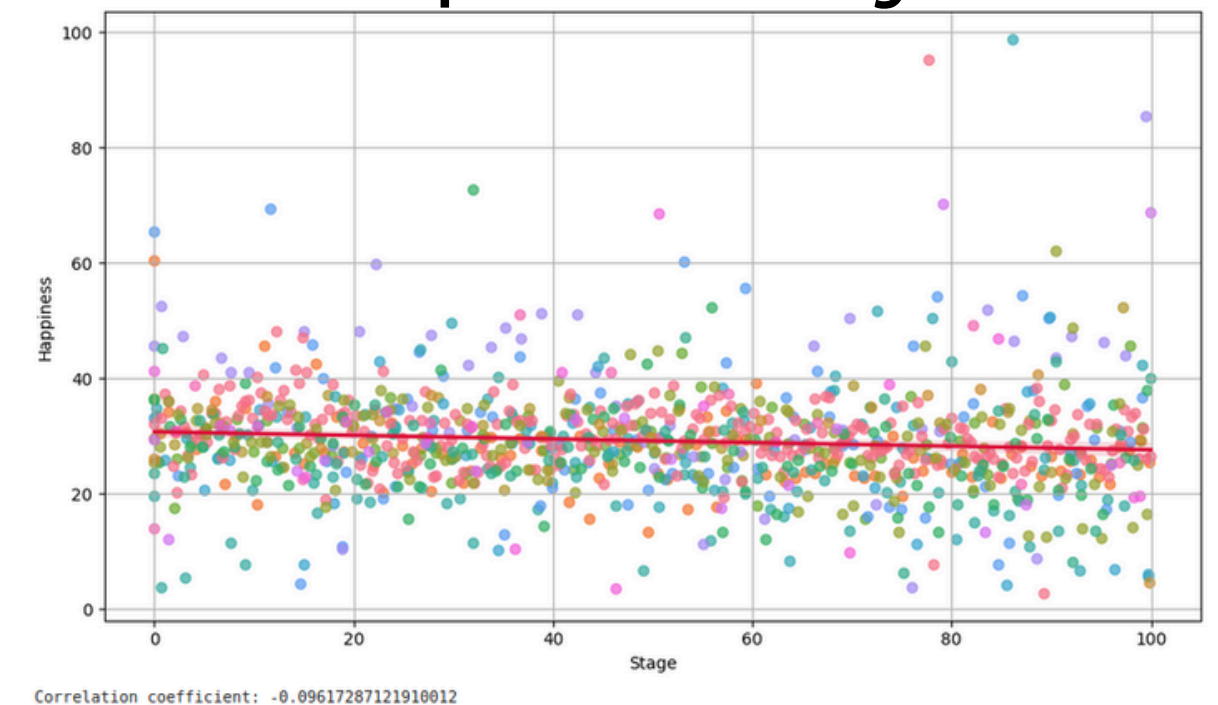
...**attention** grows?



...more **contacts** participate?



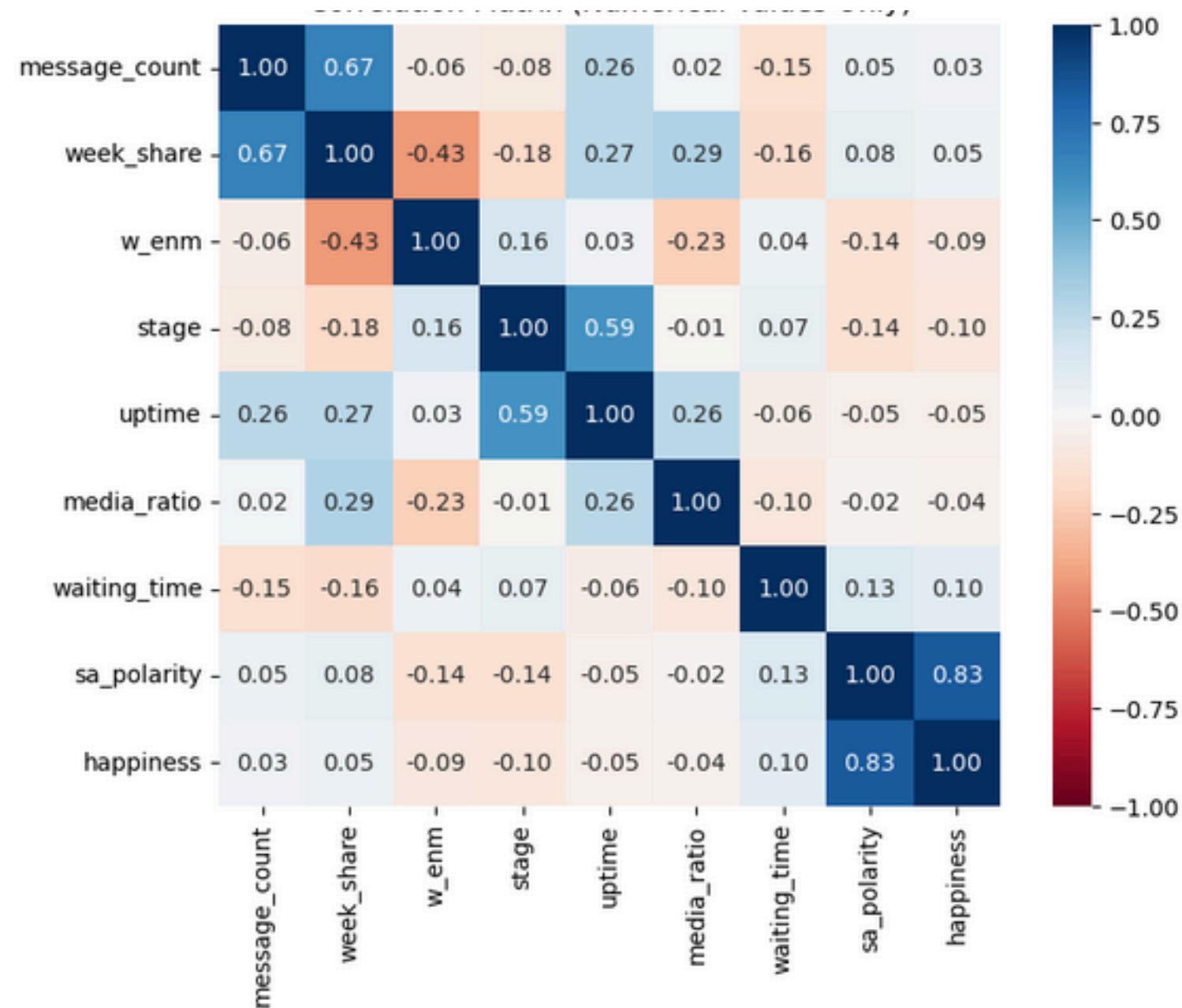
...**time** passes by?



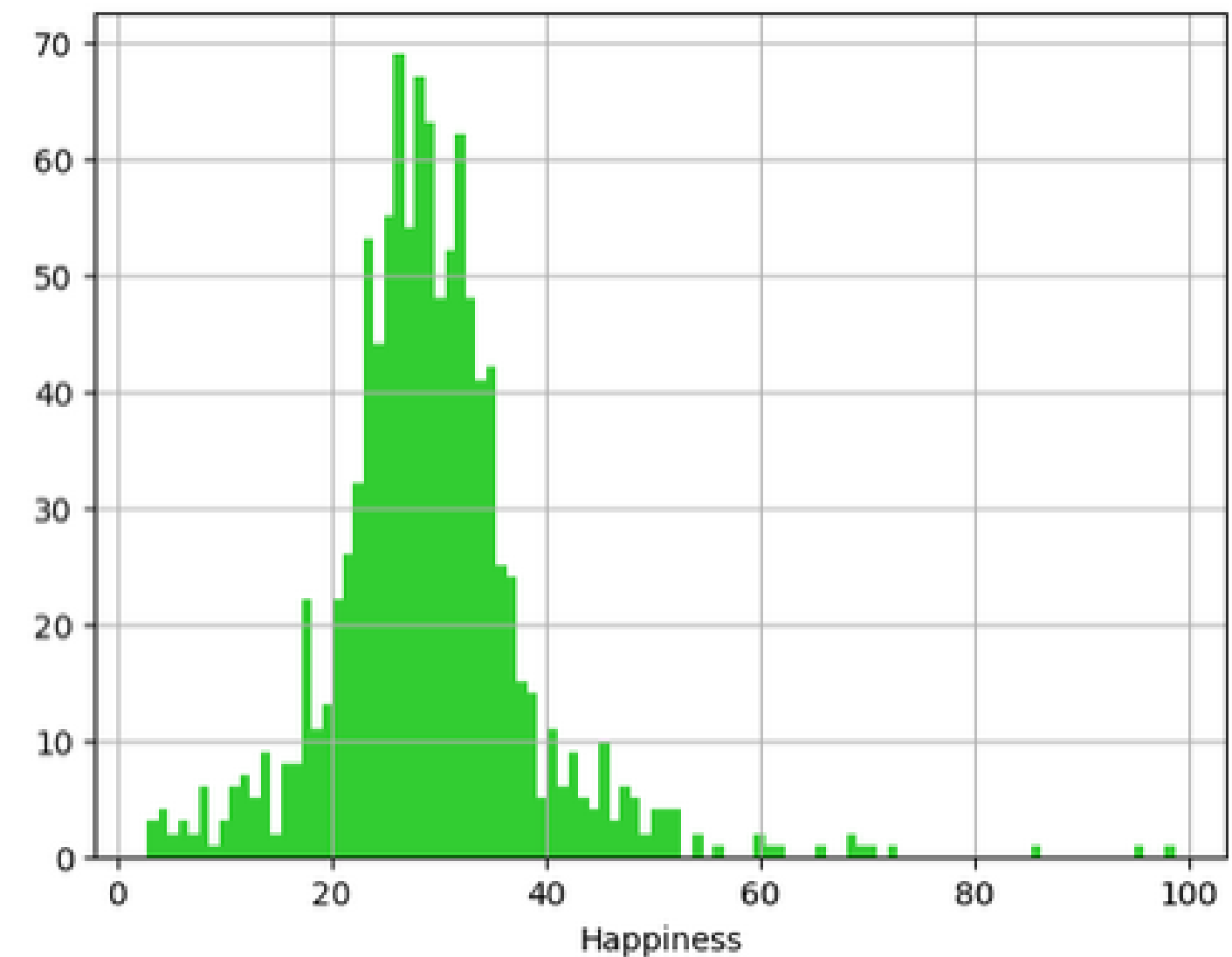
very low correlations: under 0.1
Is *happiness* a good indicator?

NO, really. What **happens** with happiness?!

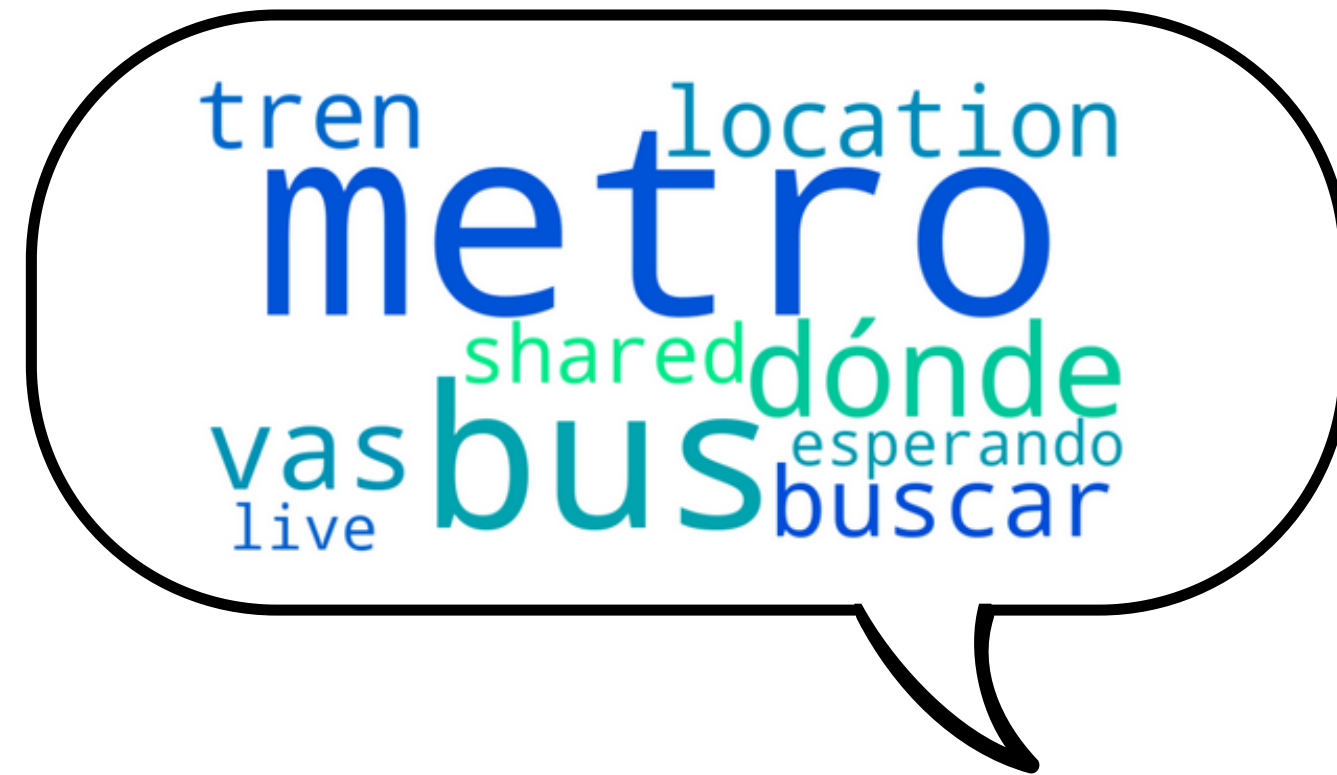
correlation matrix



happiness distribution



BERTopic topic analysis



Conclusions

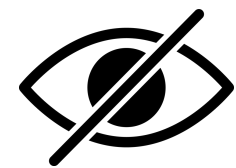


Do all interactions follow the same **overall trend**?



Do one contact activity **affect other** contacts?

Challenges



Private nature of the data



Locally run tools only



Relatively small dataset



questions and (maybe) answers

thank you **all** for listening ❤️

github.com/juanjomontero

linkedin.com/in/juanjosemontero