

Introduction

Communication is increasingly going through the world wide web, instead of the traditional analogue forms. Who can for instance remember the last time, they picked up a pen, and wrote a letter to someone, instead of simply going to WhatsApp, Facebook Messenger, Outlook, or something completely different? In this assignment, I have chosen to focus on my messages across Facebook Messenger from when I first started using Facebook, June 2008, and until present day.

This I did in order to see, if there was any particular pattern, from which I could extract some knowledge of the platform, the use of it, or maybe about myself.

Creating the data and graphs

In order to investigate my messages of Facebook, I thought that I firstly had to retrieve my Facebook Messenger "history". This is fortunately quite easy due to GDPR, where I could simply download all my messages ever sent from Facebook.

After receiving this 2.5 GB large .zip-file, I unpacked it, and began looking at the structures. From here I could create a script in Python, which went through all the messages, filtered them, and turned them into a CSV-file with only two parameters; date and amount of messages sent.

Python is great for these types of task, since it is a light-weight program, which rapidly can go through all the messages. Furthermore, it is great at accessing JSON files, which your Facebook data is, if you do not choose HTML format.

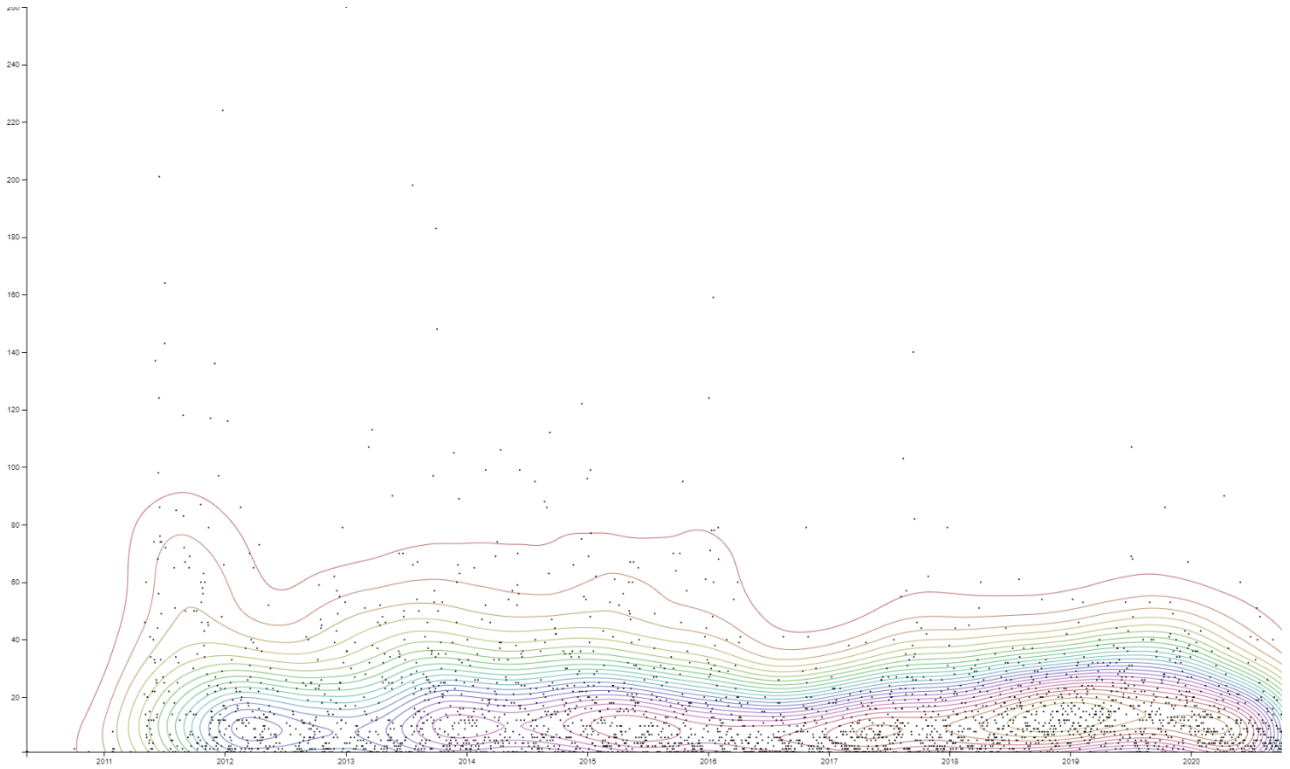
After loading the data into the Python script, I looped through all the messages, and cut away all messages that was not from me. Secondly, I filtered away message that were automatically generated, such as "You can know contact X" and so on.

Afterwards I created a dictionary in Python, which is basically just an object with keys and values, in which I generated the data while looping through the messages.

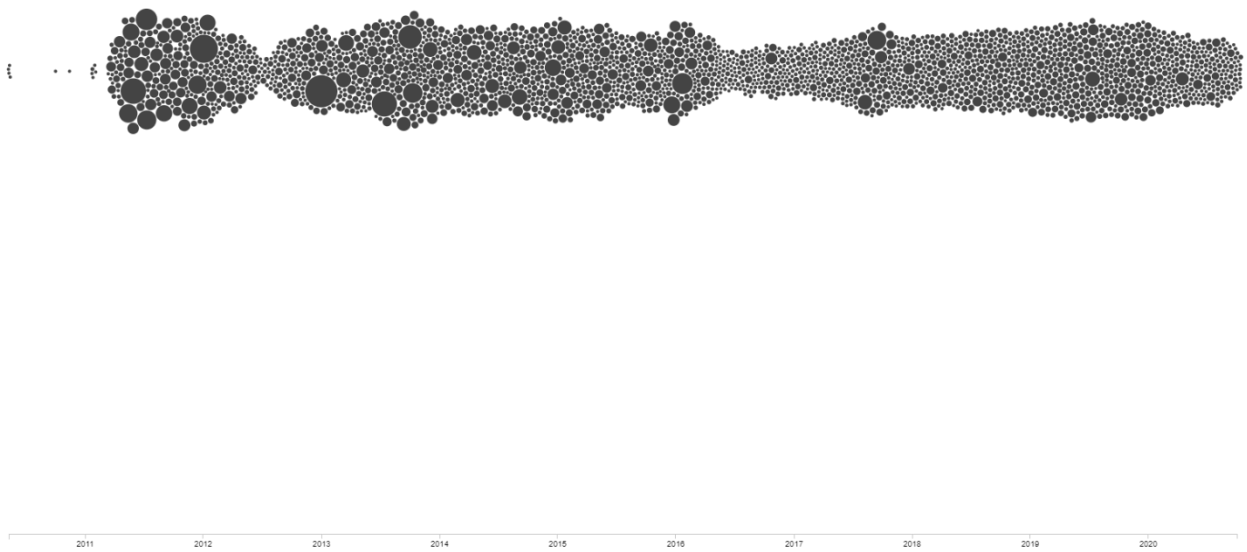
After the dictionary had been completed, I wrote the data to a .txt file, written in the form "date, amount of messages sent".

After this file had been generated, I was free to select a graph, which could display these data, so it was easier to perceive and analyze them.

Therefore I went to <https://app.rawgraphs.io/>, where you can generate graphs based on CSV files. I chose the two found on the next page (and attached), since I found they both displayed the data easily readable, but in different ways.



Graph 1



Graph 2

Analyze

The data shows, that I was most active on Facebook Messenger up until 2015, where I finished my high school diploma. The most messages was sent from 2011-2012, where I was on boarding school, which would indicate that I wrote a lot to the friends I had, who did not attend the same school. This is however contrary to the period of 2015-2018, where I was working abroad. This period was where I sent the least amount of messages, despite only being in Denmark approx. 5 weeks a year. This could be explained with, that in that period I did not have the need to organize gatherings in the weekend, or simply just different periods of life, where I was not so focused on being in contact with my network at all times.

It is however funny, to see how the amount of messages changes according to period of my life, where I can pinpoint events of life to the graph, and thinking about how it comes, that it changes the graphic as it does.

Implications

This way of using data provided by Facebook, shows that the data is not only for companies to either exploit or sell, but can also be used to provide beautiful graphs, which serves to learn about yourself. The sheer way that this data is portrayed, as a company capital, is in itself an issue, as Kreisberg and Acker states. Why do we let these multibillion companies and their complex algorithms to undertake the task of not only controlling the data itself, but also its future? This has even more implications, as this data is part of the surveillance economy (Zuboff 2015), making me a dividual (Cheney-Lippold 2017). The critical questions are then: "How does these data represent me?" and "How can they be exploited?". This is at least the thought process through a critical lens. Through a more exploratory lens, you could ask, why I would have to go through all these steps, in order to visualize how messages, I send? Then we go right back to letting the companies process our data, sorting and labeling it for algorithmic rather than exploratory use, making it obvious, that the data is mainly fuel to their engine, and not for the users.

As Annette Markham would say: The biggest accomplishment by modern companies, to lead users to believe, that data is just a waste product, and that it should be "shared".

This point is also evident in Nanna Thylstrup's *The Politics of Mass Digitization* (2018), where it is pointed out how much power these companies hold.

An example of these powers, were also evident in class, where Pablo Gonzales showed a network of his Facebook friends, with data he had retrieved prior to 2016, where Facebook shut down the API, from which users could get data about themselves. It is, when you think about it, absurd, that users are prohibited access to their own data. And not only that, it is extremely difficult (and against Facebook Terms of Use), to retrieve them anyways (through automization).

Why should it be this difficult and even against terms of conditions, to explore *our own data*?

Ethical considerations

In this research I have excluded all names, and only used my own data, since I have not retrieved permission from all persons I have written to. This is crucial due to situatedness (Haraway 1988), since they are very well aware that the data are to be seen and used by me and Facebook, but have not willingly accepted the setting of which I portray the data in this paper.

Furthermore, I have only used data provided to me by Facebook, which have been given to me within the boundaries of Facebook's Terms of Use, since it is a grey area to tread, and putting my account in risk of being disabled. This just further underlines the power these companies hold, since I am not willing to put my account on the line, meaning Facebook has created a great "lock in" mechanism for its users.

Bibliography

Thylstrup, Nanna Bonde. 2018. *The Politics of Mass Digitization*.
Cambridge, MA: The MIT Press.

Acker, Amelia, and Kreisberg, Adam. 2020. "Social Media Data Archives in an API-Driven World." *Archival Science* 20 (2): 105–23.