

Digital Shadow - Report

The Tech

I have already written about my struggles around trying to get the dataset together. But now I have a script which can let anybody pull up their Facebook dataset. The script also cleans, aggregates and organizes the data for analysis and visualization. I will shortly contribute that to the R repo for this subject and my GitHub account.

There was a lot of learning around ggplot and its many knobs, tools and short-comings. I was genuinely impressed that you can now customize every single element of the plots separately. I first started out assuming I will use ggplot to design all the elements of the plot but quickly realized how annoying it is to do so. The idea of importing transparent background plots to Illustrator so as to have a homogenous background and labels was an important step. I also learned transparent background images are called vectors and you can find them all over the internet.

The Design

As a not so creative individual I was more scared than unwilling to design using Illustrator. However, once I understood that the basic principles behind element, object and color manipulation, I was able to unlock its potential. I quickly realized the flexibility and power of using it over ggplot for designing plots. I ended up using ggplot to just create naked plots, importing them into Illustrator for the design.

The Narrative

I identified the three post common features of Facebook - Comments, Likes and Timeline.

Comments

I wanted to understand how comments differed based on time of post and type of post. The bar plot kind of tells the story that people have reduced commenting on my posts over the years. The polar plot describes that I generally get more comments on Statutes.

Likes

I created the wordcloud to figure out if my closest friends like and comment the most on posts. The idea around a wordcloud is the font size differs depending on the frequency. So, the bigger the name in the 'like' shaped wordcloud the more interaction he/she had with my posts. As soon as I looked at my dataset I could see a clear lift in my Facebook usage and likes/comments from year 2013. The horizontal bar plot highlights this.

Timeline

I designed a scatterplot of posts by likes to create the impression of a timeline. I wasn't sure if I will be able to highlight the major life events at first but fortunately I was able to. The idea here was to visualize and infer how different type of posts was getting likes over the years. Also, I wanted to see if I could highlight the type of post I used to announce the major milestones. A hack around this was to cutoff the 1991 birth post because it was an outlier in the plot. I manually inserted a separate variable to capture this.

Final Thoughts

This has been a really tough, taxing and challenging project but I have thoroughly enjoyed and learnt through the process. This was my first ever design project so a pat on the back there. I am pretty happy with the outcome.