Sentiment Analysis: Based on Instagram Comments

Prepared by

Olaoluwasubomi

i@olaoluwasubomi.com

November 2021

Introduction

Fred John Philip Gibson, known professionally as Fred Again or simply Fred, is a British record producer, singer, songwriter, multi-instrumentalist, and DJ.

Ideology: I was casually scrolling on Youtube As I always do when I'm eating and came across this Gem called FredAgain and he was doing a live performance for the boiler room. And I know Boiler room videos are always 1 hr plus and this was the first boiler room video I can say I watched from the beginning to the end and my conclusion thoughts were just…"Wow" So immediately I started to listen to more of his songs, Follow him on social media, and listen to previous and present works of his. So a day came by and I saw he posted on his Instagram 3 Samples of his soon-to-come album and asked in the caption Which does he want us to drop first between 1,2 and 3 of the samples he provided on the post and me being the Data Scientist I am... I told myself I would Predict what song is he going to release next with the help of programming

With the aid of a tool called Phantom Buster and my personal Instagram session cookie, I was able to scrape 700 of the post's 15,9k comments as the first thing I did. (Yes, I agree.)

I transferred them from the CSV format they were in into an excel spreadsheet. and the data that was gathered appeared somewhat like the attached image below

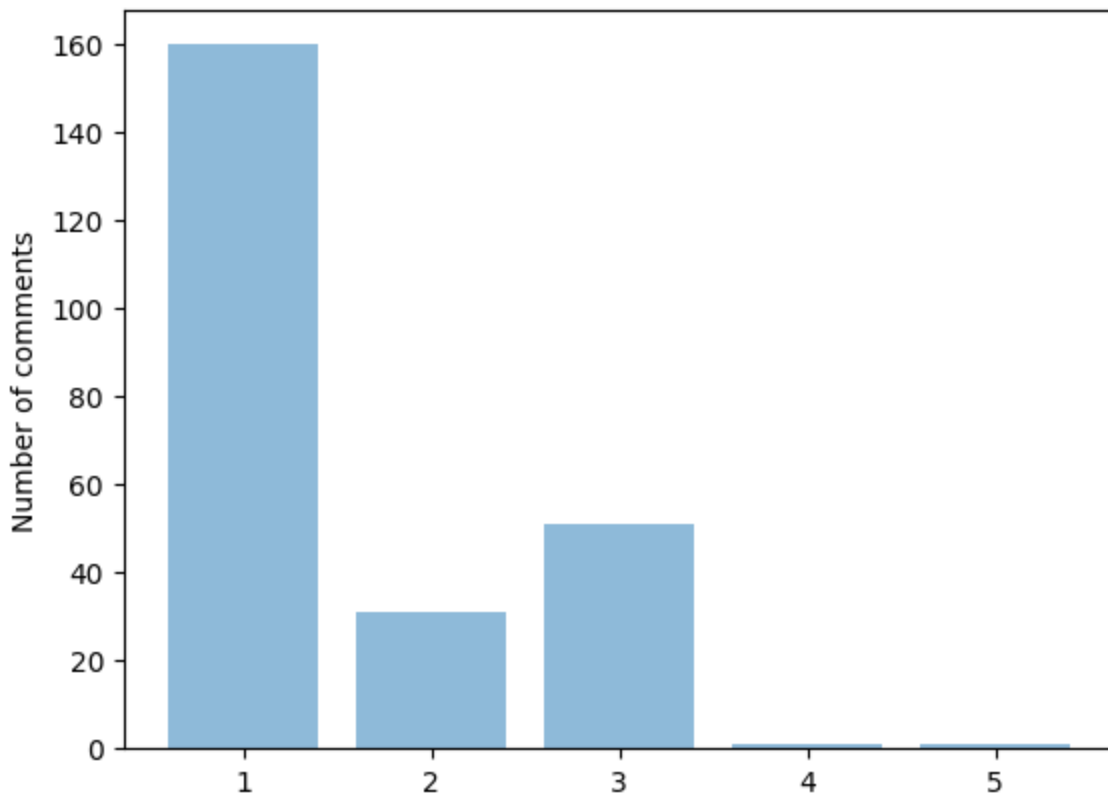|	A	B	C	D	E	F	G	H	I
1	profilePictureU	username	profileUrl	comment	likeCount	replyCount	commentDate	commentId	ownerId
2	https://sconten	dontbebilly	https://www.ir	3,ù§Ôⵠè	844	8	2022-10-07T1	17,857,227,84	306,534,230
3	https://sconten	ovyonthedrum	https://www.ir	2	27	0	2022-10-08T0	17,958,249,36	278,734,428
4	https://sconten	delilahmontag	https://www.ir	1!!!!!!!	3,628	29	2022-10-07T1	17,976,403,72	196,126,058
5	https://sconten	wilso_	https://www.ir	1!!	102	0	2022-10-07T1	17,939,577,15	13,607,711
6	https://sconten	lanostri	https://www.ir	1 NEED IT	440	0	2022-10-07T1	18,218,591,21	212,483,763
7	https://sconten	pedrolsousa5	https://www.ir	Trust me, go f	150	0	2022-10-07T1	17,987,902,21	234,571,703
8	https://sconten	clarebearita	https://www.ir	1 no question	632	0	2022-10-07T1	17,941,687,22	209,069,579
9	https://sconten	lnlhntr	https://www.ir	111111111111	311	0	2022-10-07T1	17,979,261,52	1,556,290,02
10	https://sconten	effektsvk	https://www.ir	Please 111111	918	7	2022-10-07T1	17,954,925,55	274,211,841
11	https://sconten	gracegruber	https://www.ir	fred you've be	551	4	2022-10-07T1	17,941,016,91	211,517,760
12	https://sconten	laurenz_ham	https://www.ir	Fuck please 1	92	3	2022-10-07T1	18,034,704,31	855,252,273
13	https://sconten	jack_alam10	https://www.ir	It has always t	55	0	2022-10-07T1	18,044,176,00	215,232,388
14	https://sconten	pbergen	https://www.ir	I need them al	9	0	2022-10-07T1	17,852,301,58	43,358,325
15	https://sconten	frankk.png	https://www.ir	1 PLEASEEE	291	1	2022-10-07T1	18,032,317,35	45,038,730,0
16	https://sconten	bobbysantoni	https://www.ir	1 please 1 plea	20	0	2022-10-07T1	17,976,909,72	261,580,760
17	https://sconten	chediak.club	https://www.ir	1 for the love	213	0	2022-10-07T1	17,968,221,67	15,722,229
18	https://sconten	meganbreyer	https://www.ir	1 pleaseeeee	153	0	2022-10-07T1	17,953,354,04	6,205,321

The most important on the table would be username, comment, and likecount

As we can see from the table, the comment column contains a lot of unfiltered data, such as Latin letters, the letter!, and other content that is not what we want.

I made the decision to solely use the data from the original question he asked—"Pick amongst 1, 2, and 3"—for this project. I therefore cleaned the data so that it only displayed the numbers 1, 2, and 3, and in order to improve this project in the future, I would like to suggest that the likecount be given some sort of weight as not everyone would want to comment instead of just like what they want.
In order to determine which song Fredagain would release next based on the comment data, I cleaned the comments so that they only contained the numbers 1, 2, and 3. The computer then calculated how many 1s, 2s, and 3s were in the data and created a graph for the entire project.

Number of comments for each number 1-5

244 Comments in total
FRED AGAIN IS DROPPING NUMBER 1 CALLED PULL ME OUT OF THIS

The good thing about this project is that it can be modified to anything different because it can be used for voting, polls, and so many other things in addition to Instagram comments from artists.

Source code