

Megha Mansuria
Prof. Cheryl Dugas
CS 581 WS
9 December 2020

Assignment 9: Analyze Social Network Data

PURPOSE

The program `analysis.py` retrieves data from an Internet Core Trends survey conducted by the Pew Research Center in January/February 2019. The purpose of this program is to organize the responses by the age groups that correspond to a respective generation: Generation Z, Millennials, Generation X, Baby Boomers, and the Silent Generation. Using the generations, I wanted to show the usage of two specific social media sites: Instagram and Facebook. I chose these two social media platforms because they are both owned by Facebook, Inc. This topic interested me because I was curious to see which generations use Instagram and Facebook more frequently than others. Overall, this analysis is important to analyzing social network data to see the usage breakdowns and to see the popularity of the respective social media sites to a diverse set of age groups.

INPUT

To run the program in a terminal window, enter `python3 analysis.py`.

```
TERMINAL  PROBLEMS  OUTPUT  DEBUG CONSOLE
meghamansuria@Meghas-MacBook-Pro-2 HW9 % python3 analysis.py
```

OUTPUT

There are two outputs produced by the program: (1) a written, tabular output printed to the console and (2) a figure of multiple pie charts. The printed output contains data of Instagram and Facebook usage for each generational group; the usage is separated by daily, weekly, rarely, and unknown.

```
TERMINAL  PROBLEMS  OUTPUT  DEBUG CONSOLE
meghamansuria@Meghas-MacBook-Pro-2 HW9 % python3 analysis.py

Instagram and Facebook usage by each generation:
~~~~~

Generation Z total participants: 135
      Instagram      Facebook
Total:      98      103
Daily:      77      75
Weekly:     15      22
Rarely:      6       6
Unknown:     0       0
```

Millennial total participants: 350

	Instagram	Facebook
Total:	187	279
Daily:	120	214
Weekly:	49	54
Rarely:	18	11
Unknown:	0	0

Generation X total participants: 264

	Instagram	Facebook
Total:	89	203
Daily:	50	147
Weekly:	26	42
Rarely:	13	13
Unknown:	0	1

Baby Boomer total participants: 477

	Instagram	Facebook
Total:	98	322
Daily:	36	220
Weekly:	45	80
Rarely:	17	19
Unknown:	0	3

Silent Generation total participants: 146

	Instagram	Facebook
Total:	19	82
Daily:	11	53
Weekly:	6	21
Rarely:	2	7
Unknown:	0	1

meghamansuria@Meghas-MacBook-Pro-2 HW9 %

WHAT THE PROGRAM DOES

When the program runs, I initially have it read the `Pew_Survey.csv` file. This file contains a list of responses from the survey and the program specifically looks at the 'age', 'web1b', 'web1c', 'sns2b', and 'sns2c' columns. First, each generation's respective function looks for responses from the given age range and appends the rows to a generation list. Then, the program will find and create a new list of rows that answered yes to using Instagram or Facebook in their respective for-loops. With this new list, I find the number of people who use social media daily, weekly, rarely, and unknown with several counters. For my analysis, I accepted daily as 'Several times a day' and 'About once a day'; weekly as 'A few times a week' and 'Every few weeks';

rarely as 'Less often'; and unknown as 'Don't know' and 'Refused'. At the end, I print out the outcomes of the Instagram and Facebook usage by different age groups to the console. Finally, I have 10 pie charts that appear, showing the usage percentages of Instagram and Facebook users for each generation.

RESULTS

After getting the written and visual outputs, I noticed that as the generations get older, there is a growing gap between total users of Instagram and Facebook, where the number of total Instagram users continuously decreases. Also, while Generation Z has a lower number of Instagram users (98) to Facebook users (103), more Gen Z participants use Instagram daily (78.6%) than Facebook (72.8%). Another comparison to note is how Baby Boomers use both social media sites. There are a total of 98 Instagram users and 322 Facebook users. Only 36.7% of total Baby Boomer Instagram users utilize it daily; meanwhile, 68.3% of total Baby Boomer Facebook users utilize the site daily. I was not surprised that while Instagram and Facebook have an approximately similar number of Gen Z users, that Instagram was more popular for daily usage. Similarly, my hypothesis of there being more Facebook users than Instagram users from the Baby Boomer generation was accurate. More analyses and comparisons can be seen for the Millennials, Generation Y, and the Silent Generation participants.

ADDITIONAL INFORMATION

To first determine how to define each generation's age group, I looked for a site that provided an estimate for birth years. (Source: <https://genhq.com/faq-info-about-generations/>) I use this link to define the years to determine what ages correspond to each generation. Additionally, one point to note is that while a percentage of usage may be larger, it may not necessarily mean it's the most popular social media site. For instance, in the Silent Generation, there are 11 out of 19 participants that use Instagram daily (57.9%); however, there are also 53 out of 82 participants that use Facebook daily (64.6%). These two percentages seem pretty similar for a difference of 63 total users. This was just an observation I had noticed and wanted to clarify that the percentages are of the usages (daily, weekly, rarely, unknown) to the total number of the respective social media's users in each generation.