

Social Media Use and Attitudes of Teens

Who Don't Fit In

1. Dataset Presentation

Dataset: The dataset used is the result of a survey carried out by the Pew Research Center called “Teens and Tech Survey 2018”. It seeks to understand US teenager’s use and attitude towards social media. Exactly 1,000 teens between 13 and 17 years old were surveyed.

Link to the dataset: <https://www.pewresearch.org/internet/dataset/teens-and-tech-survey-2018/>

The original dataset contained 173 columns representing a single question, each containing a teen’s answer. Below are the original dataset’s columns’ names:

```
> colnames(data)
[1] "1..CASEID" "SURV_LANG" "FITIN" "FRIEND1" "FRIEND2" "FRIEND3" "FRIEND4_1" "FRIEND4_2" "FRIEND4_3" "FRIEND4_4" "FRIEND4_5" "FRIEND4_6" "FRIEND5"
[14] "FRIEND6_1" "FRIEND6_2" "FRIEND6_3" "FRIEND6_4" "FRIEND6_5" "FRIEND6_6" "FRIEND6_7" "DEVICEA" "DEVICEB" "DEVICEC" "DEVIDE" "HOMWORKA" "HOMWORKB"
[27] "HOMWORKC" "INTREQ" "GAMING" "SNS1_1" "SNS1_2" "SNS1_3" "SNS1_4" "SNS1_5" "SNS1_6" "SNS1_7" "SNS1_8" "SNS2" "SNS3"
[40] "SOCIA_GOOD_1" "SOCIA_GOOD_2" "SOCIA_GOOD_3" "SOCIA_GOOD_4" "SOCIA_GOOD_5" "SOCIA_GOOD_6" "SOCIA_GOOD_7" "SOCIA_BAD_1" "SOCIA_BAD_2" "SOCIA_BAD_3" "SOCIA_BAD_4" "SOCIA_BAD_5" "SOCIA_BAD_6"
[53] "SOCIA_BAD_7" "SOCIA_OTHER" "SOCIA_DK_REF" "POST1A" "POST1B" "POST1C" "POST1D" "POST1E" "POST2_1" "POST2_2" "POST2_3" "POST2_4" "POST2_5"
[66] "POST2_6" "POST2_7" "POST2_8" "SOC2POSA" "SOC2POSB" "SOC2POSC" "SOC2POSD" "SOC2NEGA" "SOC2NEGB" "SOC2NEG" "SOC2NEG2" "SOC4A" "SOC4B"
[79] "SOC4C" "SOC4D" "SOC5A" "SOC5B" "SOC5C" "SOC6" "SOC7_1" "SOC7_2" "SOC7_3" "SOC7_4" "SOC7_5" "SOC7_6" "SOC7_7"
[92] "SOCXPB" "SOCXPC" "SOCXPD" "WORRYA" "WORRYB" "WORRYC" "LIMITA" "LIMITB" "LIMITC" "CELL1_1" "CELL1_2" "CELL1_3" "CELL1_4"
[105] "CELL1_5" "CELL1_6" "CELL2A" "CELL2B" "CELL2C" "CELL3A" "CELL3B" "CELL3C" "DISTRACT" "GROUP1" "GROUP2_1" "GROUP2_2" "GROUP2_3"
[118] "GROUP2_4" "GROUP2_5" "GROUP2_6" "GROUP2_7" "GROUP2_8" "GROUP2_9" "GROUP2_10" "GROUP2_11" "GROUP3A" "GROUP3B" "GROUP3C" "GROUP3D"
[131] "OH1A" "OH1B" "OH1C" "OH1D" "OH2A" "OH2B" "OH2C" "OH2D" "OH2E" "OH2F" "OH3_1" "OH3_2" "OH3_3"
[144] "OH3_4" "OH3_5" "OH3_6" "OH3_7" "OH3_8" "OH3_9" "OH3_10" "OH3_11" "OH3_12" "OH3_13" "OH3_14" "OH3_15" "OH3_16"
[157] "RACETHNICITY" "HOME_TYPE" "HOUSING" "INCOME" "INTERNET" "PHONESERVICE" "METRO" "REGION4" "HHSIZE" "HH01" "HH02" "HH03" "HH04"
[170] "HH180V" "CO_DATE" "DURATION" "SURV_MODE" "MODE_END" "DEVICE" "METRO" "REGION4" "HHSIZE" "HH01" "HH02" "HH03" "HH04" "HH05" "HH06" "HH07" "HH08" "HH09" "HH10" "HH11" "HH12" "HH13" "HH14" "HH15" "HH16" "HH17" "HH18" "HH19" "HH20" "HH21" "HH22" "HH23" "HH24" "HH25" "HH26" "HH27" "HH28" "HH29" "HH30" "HH31" "HH32" "HH33" "HH34" "HH35" "HH36" "HH37" "HH38" "HH39" "HH40" "HH41" "HH42" "HH43" "HH44" "HH45" "HH46" "HH47" "HH48" "HH49" "HH50" "HH51" "HH52" "HH53" "HH54" "HH55" "HH56" "HH57" "HH58" "HH59" "HH60" "HH61" "HH62" "HH63" "HH64" "HH65" "HH66" "HH67" "HH68" "HH69" "HH70" "HH71" "HH72" "HH73" "HH74" "HH75" "HH76" "HH77" "HH78" "HH79" "HH80" "HH81" "HH82" "HH83" "HH84" "HH85" "HH86" "HH87" "HH88" "HH89" "HH90" "HH91" "HH92" "HH93" "HH94" "HH95" "HH96" "HH97" "HH98" "HH99" "HH100"
```

Choice of questions: For this project I attempted to focus on understanding how teens which did not fit in used social media and its influence on their well-being. I decided to keep only 12 questions listed below and include only the answers of teenagers who had answered “I tend to stand out” to the question ‘Fitin’ on how well they thought they fit in with others:

```
> colnames(data)
[1] "FITIN" "SOC1" "SOC2NEGA" "SOC4A" "SOC4B" "SOC4C" "SOC4D" "GROUP1" "GROUP3A"
[10] "GROUP3B" "GROUP3C" "GROUP3D"
```

1. [FITIN] In general, which of the following statements comes closest to describing how you see yourself compared with other people your age where you live?
2. [SOC1] Overall, what effect would you say social media has had on people your age?
3. [SOC2NEGA] (a-d) In general, does what you see on social media make you feel...
4. [SOC4] (a-d) How social media makes them feel...
5. [GROUP1] How often, if ever, do you spend time in any groups or forums online (including on blogs, social media, or online gaming forums)?
6. [GROUP3] (a-d) Thinking about online groups where you spend time, how much of a role, if any, have they had in...

Note: To fit the conditions of certain questions which require certain answers to certain questions, I decided to remove all participants who had answered questions in a manner that was incompatible with the questions that I decided to keep.

2. Multiple Correspondence Analysis:

```
> #percentage of variability explained
> head(eig.val)
```

	eigenvalue	percentage of variance	cumulative percentage of variance
dim 1	0.25844359	17.229573	17.22957
dim 2	0.21016341	14.010894	31.24047
dim 3	0.16127458	10.751639	41.99211
dim 4	0.09712141	6.474761	48.46687
dim 5	0.08619226	5.746151	54.21302
dim 6	0.08057617	5.371745	59.58476

The first 6 dimensions of MCA is only able to explain ~60% of the variability in the dataset. Since there are 12 columns, the average axis should account for approximately:

$$\frac{1}{\#columns - 1} = \frac{1}{12} = 8.333 \dots \%$$

in terms of the 12 columns.

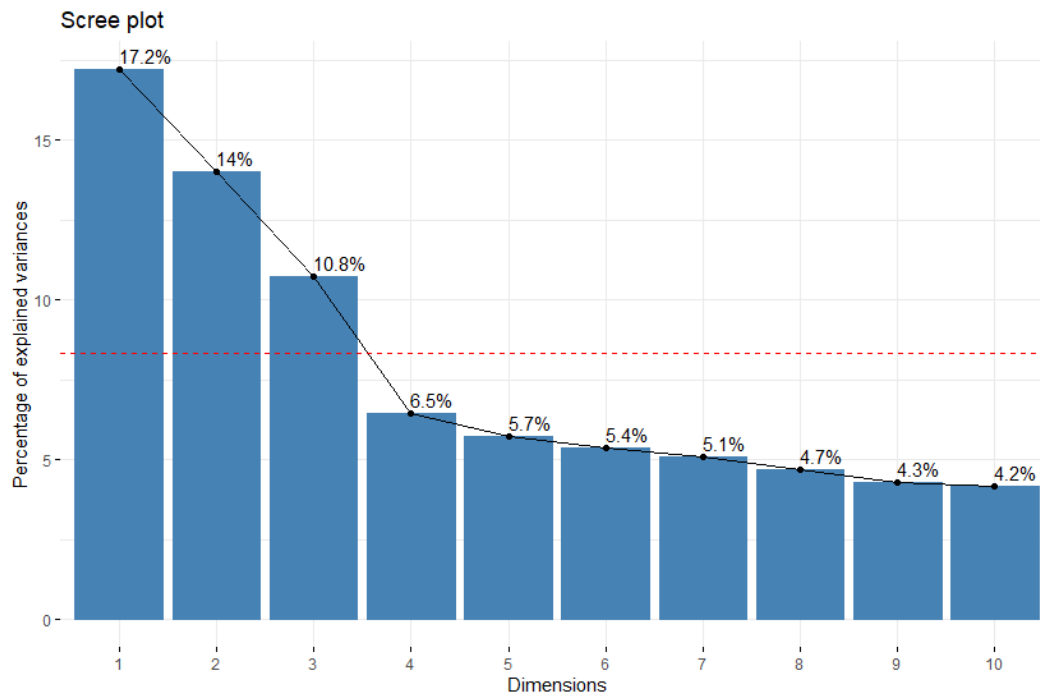
Since there are 277 rows in the reduced dataset, the average axis should account for approximately

$$\frac{1}{\#rows - 1} = \frac{1}{276} = 0.362 \dots \%$$

in terms of the 277 rows.

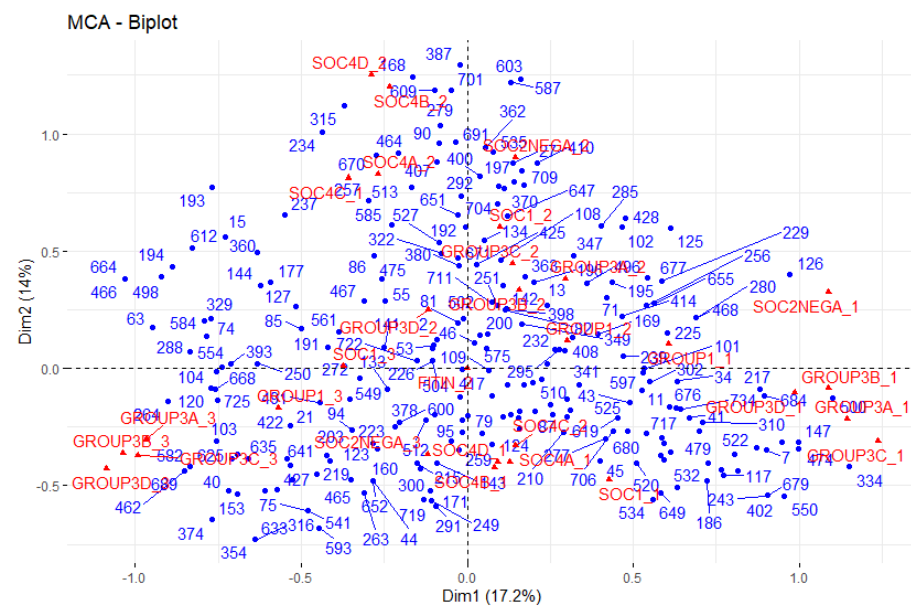
The average eigenvalue should be $\max\{0.362, 8.333\} = 8.333\%$ so we can pick dimensions 1,2 and 3 to explain the variability in the dataset. The remaining dimensions can be cast aside since they account for < 8.333% of the variability.

Running a scree plot on the percentage of variance explained by each dimension we can support our conclusion that only the first three dimensions should be kept:



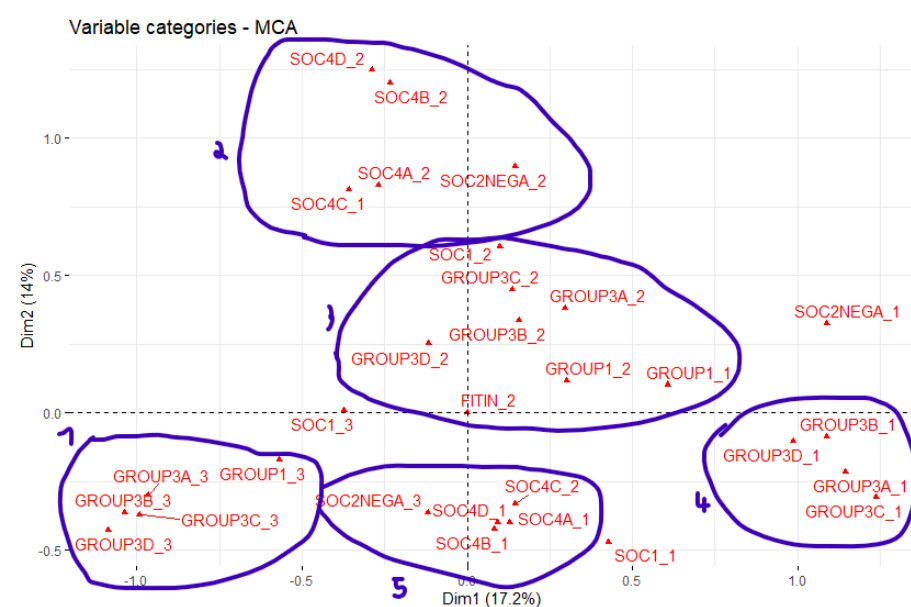
Global Trends:

Displaying a biplot of the global pattern between individuals who took part in the survey and the category variables we do not find any clear relationship. There does not seem to be any significant clusters of individuals around certain groups of variables. Individuals seem spread out evenly across both dimensions with a group of a dozen individuals (664, 194, 466, 498, 63...) whose profile seems to be in a sparsely populated area in the first two dimensions and who are far from other categories. However, since categories belong to column variables and individuals row variables, their distance is not meaningful; only the distances between individuals and the distances between categories are meaningful.



Relationship between category variables:

The following graph helps us visualize the relationship that different category variables have with each other:



While there are not any significant groupings, we can identify 4 different clusters concentrated in dimension 1 and dimension 2 (circled by yours truly in blue) in the graph.

Cluster 1: regroupes the answers to the questions:

1. GROUP3 “Thinking about online groups where you spend time, how much of a role, if any, have they had in...”
 - a. Helping you get through tough times in your life – 3 No role
 - b. Making you feel more accepted – 3 No role
 - c. Helping you figure out how to feel about important issues – 3 No role
2. GROUP1 “How often, if ever, do you spend time in any groups or forums online (including on blogs, social media, or online gaming forums)?” – 3 Rarely

Since I have decided to remove all respondents who confirmed they do not use social media or online forums, these respondents have not found online communities on the platforms they use. However, that these answers are clustered together makes sense. Positive answers to one question in 1, would lead to positive answers in all questions; having a community that makes you feel accepted would help a teen go through rough times and help them understand different issues going on in their lives. Furthermore, as indicated in question 2 these teens are not frequent users of social media and online forums. It would seem normal that teens who very rarely use online communities would develop close ties with other users.

Cluster 2: regroupes answers to questions:

- SOC4a “Does social media make you feel more...” - 2. Fake
- SOC4b “Does social media make you feel more...” - 2. Insecure
- SOC4c “Does social media make you feel more...” - 1. Reserved
- SOC4d “Does social media make you feel more...” - 2. Excluded
- SOC2NEGa “In general, does what you see on social media make you feel worse about your life?": 2. Yes, a little

We can see that the answers clustered in this group express negative attitudes towards the role of social media in the respondent's lives. Respondents generally feel like social media is a negative influence in their lives which promotes bad feelings. That they are all clustered together also makes sense: feeling excluded by social media can be caused by feelings of insecurity and doubt concerning the genuineness of social media posts followed by feelings that a teen should be reserved. The population focused on in this dataset were teens who did not fit in with their peers. It would thus seem normal that holding one negative attitude towards social media content would be followed by numerous other negative sentiments.

Cluster 3: regroupes questions:

1. GROUP3 “Thinking about online groups where you spend time, how much of a role, if any, have they had in...”
 - a. Helping you get through tough times in your life – 2 Minor role
 - b. Making you feel more accepted – 2 Minor role
 - c. Helping you figure out how to feel about important issues – 2 Minor role
2. GROUP1 “How often, if ever, do you spend time in any groups or forums online (including on blogs, social media, or online gaming forums)?” – 1 Often & 2 Sometimes
3. SOC1 “Overall, what effect would you say social media has had on people your age?” – 2 Mostly Negative

In this cluster are concentrated respondents who do not have strong opinions on social media and online forums. They spend either a lot of time or little time on social media or online forums. This cluster makes sense when opposed to Cluster 1. These respondents are users of online forums and social media and thus would not develop strong negative attitudes towards these platforms since they use them either regularly or occasionally. What is interesting is that even though these platforms play a little bit of a role on their well-being and they use them frequently or occasionally, they still have negative attitudes towards social media and online forums.

Cluster 4: regroups answers to questions

1. GROUP3 “Thinking about online groups where you spend time, how much of a role, if any, have they had in...”
 - a. Helping you get through tough times in your life – 1 Major role
 - b. Making you feel more accepted – 1 Major role
 - c. Helping you figure out how to feel about important issues – 1 Major role

This cluster just the opposite cluster to Cluster 1. The same reasoning that applies to Cluster 1 will apply to Cluster 4: respondents which have found online communities with which they have good interactions will help them get through hard times, make them feel accepted, help them figure out important issues and expose them to new people. These answers follow each other and to answer positively to one intuitively means you would answer positively to all questions.

Cluster 5: regroups answers to questions

- SOC4a “Does social media make you feel more...” - 1. Authentic
- SOC4b “Does social media make you feel more...” - 1. Confident
- SOC4c “Does social media make you feel more...” - 2. Outgoing
- SOC4d “Does social media make you feel more...” - 1. Included
- SOC2NEGa “In general, does what you see on social media make you feel worse about your life?”: 3. No

This group is the opposite group to Cluster 2; respondents have positive attitudes towards social media since it makes them feel good about themselves and thus it makes sense that social media doesn't make them feel worse about their own life.

Conclusion:

Many of the clusters we found seem to support intuitively reasonable facts about what type of attitudes are grouped together. Teens who have formed bonds in online communities display positive attitudes towards social media while those who have not tend to develop negative attitudes towards social media. However, all conclusions are far from definitive and are tentative at most. Firstly, I am neither an expert on teen psychology so have a limited ability to interpret the data. Secondly, the first two dimensions only explain about 31% of the variance. Thus, the groupings found by MCA in dimensions 1 and 2 are not reliable indicators of the general variability in the dataset.

