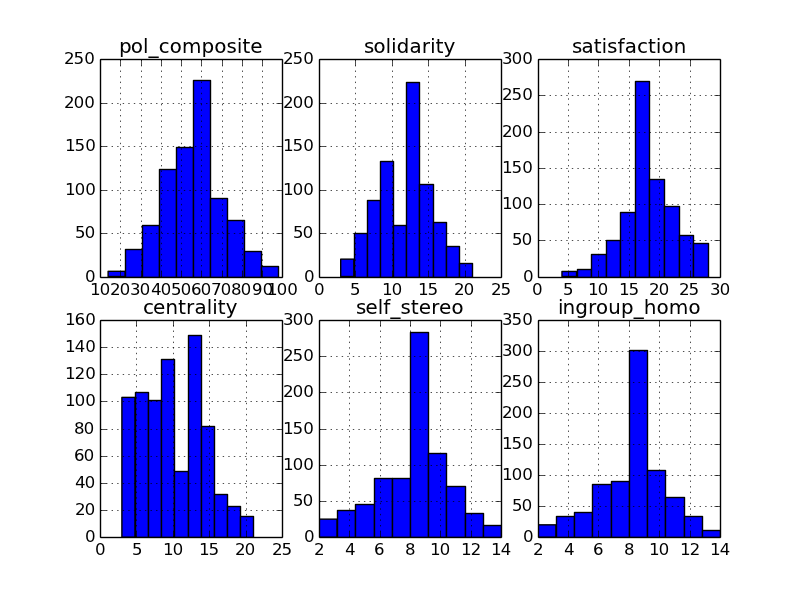
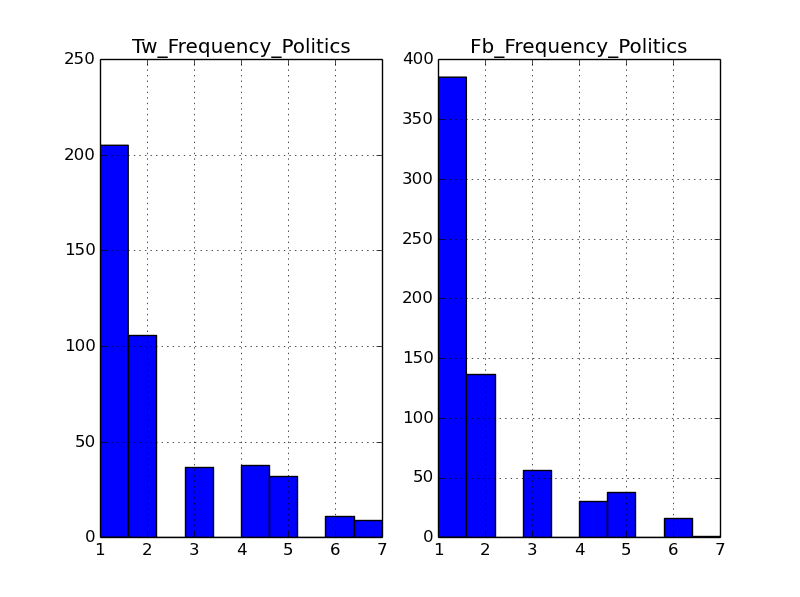
Initial Observations from Survey

**Levels of Identification with Political Party.**

(From 14 item, 5 factor scale by Leach et al)- ‘pol\_composite’= aggregated score from all 14 questions.



Distribution of responses to the question- ‘How often do you express anything related to politics on Twitter/ Facebook?’ 1=Never, 7=Constantly. [I am paraphrasing- need to go back to survey code.]



**Correlation between Levels of Identification and Stated Frequency of Expressing Politics in Social Media.**

Of all the identity subcomponents, ‘centrality’ appears to be the one most correlated with how much someone claims to express political views in social media.

CORRELATIONS OF Identity Components with claimed frequency of expression of politics on **Facebook**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Component | Composite | Solidarity | Satisfaction | Centrality | Self-Stereotyping | In-group Homogeneity |
| Pearson Corr | 0.233174 | 0.234175 | 0.201197 | **0.262479** | 0.120841 | 0.093715 |

For **Twitter:**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Sub-Component | Composite | Solidarity | Satisfaction | Centrality | Self-Stereotyping | In-group Homogeneity |
| Pearson Corr | 0.292594 | 0.289040 | 0.235932 | **0.322969** | 0.191640 | 0.150594 |

**Note** the relative patterns in correlation- in terms of which factors are stronger are exactly the same for both social networks.

Lastly, note that identity components seem to be more correlated for Twitter than Facebook (this pattern holds for every subcomponent and the composite. )

How often do you express politics on Facebook?

survey['fb\_political'].describe()

Out[83]:

count 663.000000

mean 1.870287

std 1.331040

min 1.000000

25% 1.000000

50% 1.000000

75% 2.000000

max 7.000000

dtype: float64

How often do you express politics on Twitter?

survey['tw\_political'].describe()

Out[84]:

count 438.000000

mean 2.212329

std 1.559831

min 1.000000

25% 1.000000

50% 2.000000

75% 3.000000

max 7.000000

dtype: float64

The difference between how much people report expressing their political leanings on twitter vs. facebook is significant:

t, p = scipy.stats.ttest\_ind(survey['fb\_political'].dropna(), survey['tw\_political'].dropna())

t, p

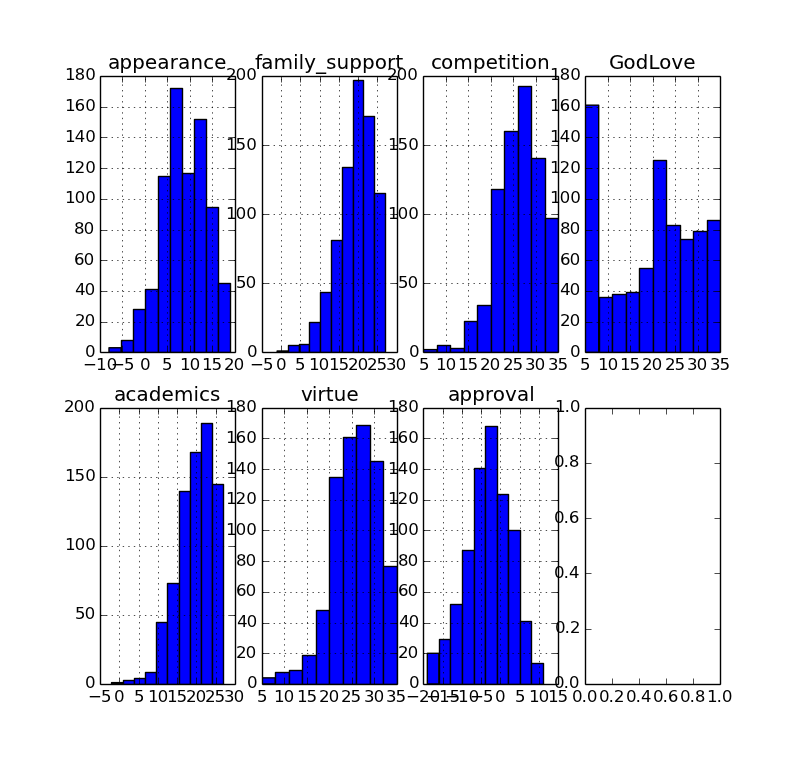
Out[94]: (array(-3.8943345927667172), 0.00010440613589120882)

This is the p-value: 0.00010440613589120882

This might just be because we phrase it in terms of frequency- people might just more frequently talk about post on Twitter than facebook, hence the likelihood of them posting more frequently on any topic on Twitter will be higher. I guess we could control for the frequency of tweets.

Generally our sample seems to claim to rarely express their political identity on Facebook or Twitter.

**Distributions for Contingencies of Self-Worth**



Comparison of Means for Expression of CSWs on Twitter and Facebook

t\_comparisons(survey,pair\_list)

The Mean for **fb\_academic** is 3.09789156627 , and the mean for **tw\_academic** is 3.81735159817 ;The p-value for this comparison is 5.48909902974e-12 ,and the t-value is -6.96901487129

The Mean for **fb\_appearance** is 2.46525679758 , and the mean for **tw\_appearance** is 2.70681818182 ;The p-value for this comparison is 0.0131106662989 ,and the t-value is -2.48478874151

The Mean for **fb\_doing** is 2.87368421053 , and the mean for **tw\_doing** is 4.23287671233 ;The p-value for this comparison is 1.38120628588e-38 ,and the t-value is -13.5080510966

The Mean for **fb\_entertain** is 3.32880844646 , and the mean for **tw\_entertain** is 4.425 ;The p-value for this comparison is 3.38198054644e-20 ,and the t-value is -9.38833753309

The Mean for **fb\_family** is 3.24471299094 , and the mean for **tw\_family** is 3.44874715262 ;The p-value for this comparison is 0.0643705560676 ,and the t-value is -1.85146661755

The Mean for **fb\_feel** is 2.38159879336 , and the mean for **tw\_feel** is 4.24772727273 ;The p-value for this comparison is 1.88622233331e-66 ,and the t-value is -18.4516476962

The Mean for **fb\_god** is 1.98340874811 , and the mean for **tw\_god** is 2.16400911162 ;The p-value for this comparison is 0.0513811622497 ,and the t-value is -1.95040838502

The Mean for **fb\_political** is 1.87028657617 , and the mean for **tw\_political** is 2.21232876712 ;The p-value for this comparison is 0.000104406135891 ,and the t-value is -3.89433459277

The Mean for **fb\_where** is 2.75825825826 , and the mean for **tw\_where** is 3.62700228833 ;The p-value for this comparison is 2.19924839042e-18 ,and the t-value is -8.90281658799

Correlation between CSW and stated frequency of expression of this CSW on Facebook and Twitter

Here are the correlations between CSW’s how much people claim to express things in social media.

Pair\_Cor(survey,pair\_list\_corr)

The correlation between fb\_academic and academics is 0.214471183582 , the t statistic is 5.64967252449 and the p-value is 1.19435945578e-08

The correlation between fb\_appearance and appearance is 0.191776447681 , the t statistic is 5.02000427926 and the p-value is 3.32743396503e-07

The correlation between fb\_family and family\_support is 0.222622227855 , the t statistic is 5.86648938889 and the p-value is 3.52109763568e-09

The correlation between fb\_god and GodLove is 0.463420587309 , the t statistic is 13.4454230012 and the p-value is 0.0

The correlation between tw\_academic and academics is 0.202288273974 , the t statistic is 4.31307151955 and the p-value is 9.96555231991e-06

The correlation between tw\_appearance and appearance is 0.222529741207 , the t statistic is 4.77698094792 and the p-value is 1.21582683321e-06

The correlation between tw\_family and family\_support is 0.170152745147 , the t statistic is 3.60960201074 and the p-value is 0.000171089378947

The correlation between tw\_god and GodLove is 0.519064128719 , the t statistic is 12.694923723 and the p-value is 0.0

Cross Platform Correlations of Stated Frequency of Expression of CSWs in Social Media

For a number of people- a few hundred- we have both Twitter and Facebook answers to comments- so we can essentially look at correlations between people expressing things across platforms- some examples:

p\_corr(survey['fb\_academic'],survey['tw\_academic'])

Out[119]: (0.42991719144796869, 12.251503560776497, 0.0)

p\_corr(survey['fb\_feel'],survey['tw\_feel'])

Out[120]: (0.35255719246135325, 9.6861632514276881, 0.0)

p\_corr(survey['fb\_where'],survey['tw\_where'])

Out[121]: (0.36904597082145862, 10.231908154290508, 0.0)

p\_corr(survey['fb\_entertain'],survey['tw\_entertain'])

Out[122]: (0.34171382416717838, 9.3481562703958705, 0.0)

p\_corr(survey['fb\_god'],survey['tw\_god'])

Out[123]: (0.61268866477643691, 19.931280673237605, 0.0)

p\_corr(survey['fb\_appearance'],survey['tw\_appearance'])

Out[124]: (0.36627315971481772, 10.112469942614938, 0.0)

p\_corr(survey['fb\_family'],survey['tw\_family'])

Out[125]: (0.47725623489801233, 13.952474321724946, 0.0)

p\_corr(survey['fb\_doing'],survey['tw\_doing'])

Out[126]: (0.20305678242915953, 5.339708044287641, 6.4025369117537423e-08)

p\_corr(survey['fb\_political'],survey['tw\_political'])

Out[127]: (0.58584640174890434, 18.585455804572558, 0.0)

Things to note- again God is a special topic, in that if people are likely to Tweet about it, they are also likely to express it on Facebook. Politics are kind of the same way. There is only a weak correlation between whether someone posts what they happen to be doing on Twitter and whether they also do this on Facebook.

[Note: still need to see if there is any testing data from me or Winter running through the experiment to check things- this will have a negligible impact on the results]

**Self Aspects by Category Label (Frequency Distribution)**

true\_selves 264

relationships 241

roles 221

public 159

situations 159

goals 155

alone 90

emotional 82

time 44

other 40

Comparisons of Frequently people claim to express different types of Self Aspects on Twitter and Facebook

For the comparison of the category value **other**, the mean for Twitter is 3.86956521739 , and the mean for Facebook is 3.0 ;The p-value for this comparison is 0.0977627401972 ,and the t-value is 1.68612974284

For the comparison of the category value **goals**, the mean for Twitter is 3.85567010309 , and the mean for Facebook is 3.06015037594 ;The p-value for this comparison is 0.000860127329017 ,and the t-value is 3.37746225612

For the comparison of the category value **time**, the mean for Twitter is 3.95652173913 , and the mean for Facebook is 3.16666666667 ;The p-value for this comparison is 0.115084342423 ,and the t-value is 1.59783227582

For the comparison of the category value **emotional**, the mean for Twitter is 3.15384615385 , and the mean for Facebook is 2.91304347826 ;The p-value for this comparison is 0.520562080283 ,and the t-value is 0.644626195414

For the comparison of the category value **roles**, the mean for Twitter is 3.95 , and the mean for Facebook is 3.48677248677 ;The p-value for this comparison is 0.0167389684772 ,and the t-value is 2.40474977503

For the comparison of the category value **situations**, the mean for Twitter is 4.13541666667 , and the mean for Facebook is 3.63970588235 ;The p-value for this comparison is 0.0476964685054 ,and the t-value is 1.9907090816

For the comparison of the category value **relationships**, the mean for Twitter is 4.19867549669 , and the mean for Facebook is 3.94634146341 ;The p-value for this comparison is 0.218978093078 ,and the t-value is 1.23143381675

For the comparison of the category value alone, the mean for Twitter is 3.0 , and the mean for Facebook is 2.25316455696 ;The p-value for this comparison is 0.0195024684404 ,and the t-value is 2.36468136124

**For the comparison of the category value true\_selves, the mean for Twitter is 4.55479452055 , and the mean for Facebook is 3.55454545455 ;The p-value for this comparison is 4.8192953773e-07 ,and the t-value is 5.12585833224**

**For some reason my t-test code is not working for the comparison of ‘public selves’- but here is what I can find out for now**

**k['Twitter'].describe()**

**Out[175]:**

**count 98.000000**

**mean 4.275510**

**std 1.870997**

**min 1.000000**

**25% 3.000000**

**50% 4.500000**

**75% 6.000000**

**max 7.000000**

**dtype: float64**

**k['Facebook'].describe()**

**Out[176]:**

**count 133.000000**

**mean 3.721805**

**std 1.904355**

**min 1.000000**

**25% 2.000000**

**50% 4.000000**

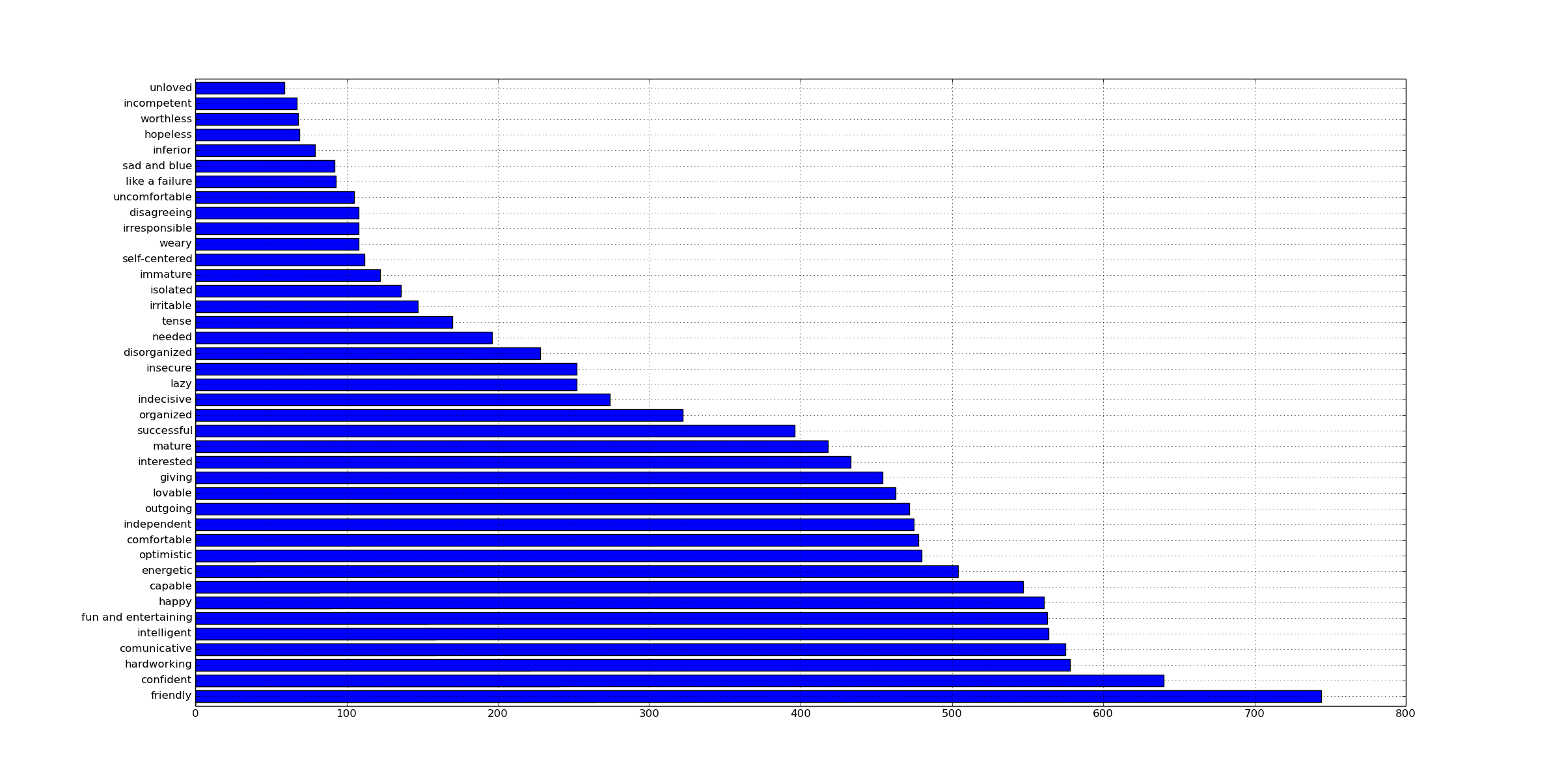
**75% 5.000000**

**max 7.000000**

**dtype: float64**

* **This comparison is probably significant. Overall, it looks like people more frequently express self-aspects on twitter.**

**Positivity Bias in valence of traits used in self aspects**

****

Other analyses to look at- reported levels of trait emotions- and the extent to which people report expressing self-aspects and other things like that in social media.

Positivity bias- is there a correlation between how positive something is and how likely they are to express it on FB or Twitter? In short, yes- there is a moderate correlation- so there might be a bit of a positivity bias.

Pair\_Cor(aspects,[('Positive','Facebook'),('Positive','Twitter')])

The correlation between Positive and Facebook is 0.335842729177 , the t statistic is 13.5207960635 and the p-value is 0.0

The correlation between Positive and Twitter is 0.286175569044 , the t statistic is 11.3257290144 and the p-value is 0.0

[This might be a little stronger for FB than twitter- which is something that we found earlier].

It’s actually pretty cool that we get this correlation because people tend to feel pretty positively about their self-aspects: The median rating is 6 out of 7.

aspects.Positive.describe()

Out[198]:

count 1440.000000

mean 5.490972

std 1.827495

min 0.000000

25% 5.000000

50% 6.000000

75% 7.000000

max 7.000000

dtype: float64

Is there a relationship between how positively people feel about a given self-aspect and the relative number of positive or negative traits present in the self-aspect?

Interestingly, the importance of the self-aspect had less of a relationship between how frequently people reported expressing a self-aspect on Twitter or Facebook:

Pair\_Cor(aspects,[('Important','Facebook'),('Important','Twitter')])

The correlation between Important and Facebook is 0.231849674288 , the t statistic is 9.03824203488 and the p-value is 0.0

The correlation between Important and Twitter is 0.173138312595 , the t statistic is 6.66624974 and the p-value is 1.86474169439e-11

Okay, other idea- we have a sense of how normative certain traits are for self-aspects- based on frequency. I could create a score for the normativity of a given self-aspect based on its popularity. My guess is that the less normative the self-aspect the less likely people would express it on either Twitter or Facebook- but maybe people would be even less likely to express it on Facebook.

Also, is there a relationship between the PANAS and how many positive or negative traits appeared in the self-aspects generated?

Pair\_Cor(aspects,[('avg\_val','Twitter'),('avg\_val','Facebook'),('avg\_val','Positive'),('avg\_val','Important')])

The correlation between avg\_val and Twitter is 0.277457540182 , the t statistic is 11.177572995 and the p-value is 0.0

The correlation between avg\_val and Facebook is 0.31088474642 , the t statistic is 12.6598084319 and the p-value is 0.0

The correlation between avg\_val and Positive is 0.731192242517 , the t statistic is 41.4852712701 and the p-value is 0.0

The correlation between avg\_val and Important is 0.523738481198 , the t statistic is 23.7953680779 and the p-value is 0.0

Tried to see if a normativity score was a better predictor- nope- it’s basically the same or a little worse (I think because it is so highly correlated with valence in the first place)

Here is the output, though:

Pair\_Cor(aspects,[('norm\_avg','Twitter'),('norm\_avg','Facebook'),('norm\_avg','Positive'),('norm\_avg','Important')])

The correlation between norm\_avg and Twitter is 0.278663284759 , the t statistic is 11.230227191 and the p-value is 0.0

The correlation between norm\_avg and Facebook is 0.305085287919 , the t statistic is 12.3991516519 and the p-value is 0.0

The correlation between norm\_avg and Positive is 0.710318210443 , the t statistic is 39.0579817553 and the p-value is 0.0

The correlation between norm\_avg and Important is 0.50373304493 , the t statistic is 22.5690509803 and the p-value is 0.0

Random Musings

Musings- it feels like Twitter is the platform where people do more EXPRESSION. We could say that Twitter Users are a self-selected group who want to do that kind of thing- maybe they systematically differ (we can’t really say for sure from our sample, because we only know when someone is both a Twitter user and a Facebook user- if they just authorize Twitter or Facebook). It should also be noted that these trends hold for our demographic- basically 18 and 19 year olds in college. However, this college demographic is extremely important and they set trends- taking the pulse with them is a good idea.

Interesting, Facebook is far more private- so why not feel comfortable to express whatever is on your mind? So, first of all, it could be a matter of people just using Twitter more primarily for expression for venting- Facebook has a lot of other features- but one would assume that these features are in some way an expression of an identity or self-aspect. Posting a picture with your girlfriend, or an article should express something about who you are. But maybe with the newsfeed option, it paradoxically causes people to produce less content than before- No, now their content is broadcast. Well, I guess we have the qualitative data, too, so we can start to code it.

Can we actually find differences in semantics from posts on twitter vs. facebook? Are certain words more common? Are there just different norms on these platforms.

Are there people who just claim to reveal more- on Twitter and Facebook- controlling for the quantity tweets or posts- are there differences in their language? There might be 25 percent of our sample or something like that- that are just deep sharers- this is kind of like the concept of a ‘me-former’, but with our data we might be able to be more explicit about the types of identities they share and how they do this.

In general it is interesting because Facebook and Twitter- at least the data we have is really concerned with presentation to an audience. It’s not about just one person- it’s about what you want to show to many friends, or in some cases with Twitter, the Internet at large. I think young people might be getting more used to living in front of people. Obviously our sample might be biased in a certain way- after all the people who are willing to give access to FB and Twitter are more likely to share things- However, much of the FB content from our users was listed as private and we took 800 people out of a possible pool of less than 4000- so even if these students are somehow different, they represent a substantial minority of college students enrolled in psychology classes- and since these classes are not for majors they represent many different types of students across the university. Since we need consent, there is really no other way to do a study like this.