MBTI Project

Data Story (EDA)

BACKGROUND

The Myers-Briggs Type Indicator (MBTI) is a well known and widely used personality inventory based on the psychological theories of Carl Gustav Jung. It is often used as a profiling tool for discovering and understanding different human personalities. It should be noted that MBTI is not a test, there are no right or wrong answers, and it does not reveal everything about oneself, additionally, it is based subjectively on the perception of each person about themselves. Nevertheless, it has been a source of personal and professional growth for many people and further studies in this field can be beneficial.

DATA

The dataset comes from Kaggle. It consists of +8600 rows with two columns. One of them is the MBTI 4 letter type of a certain person, and the other one is a collection of the last 50 things they have posted in the website PersonalityCafe. After doing some data wrangling on this dataset, here we present some insights of the Exploratory Data Analysis (EDA) process.

LEVELS OF ANALYSIS

There are three possible levels of analysis with the MBTI dataset:

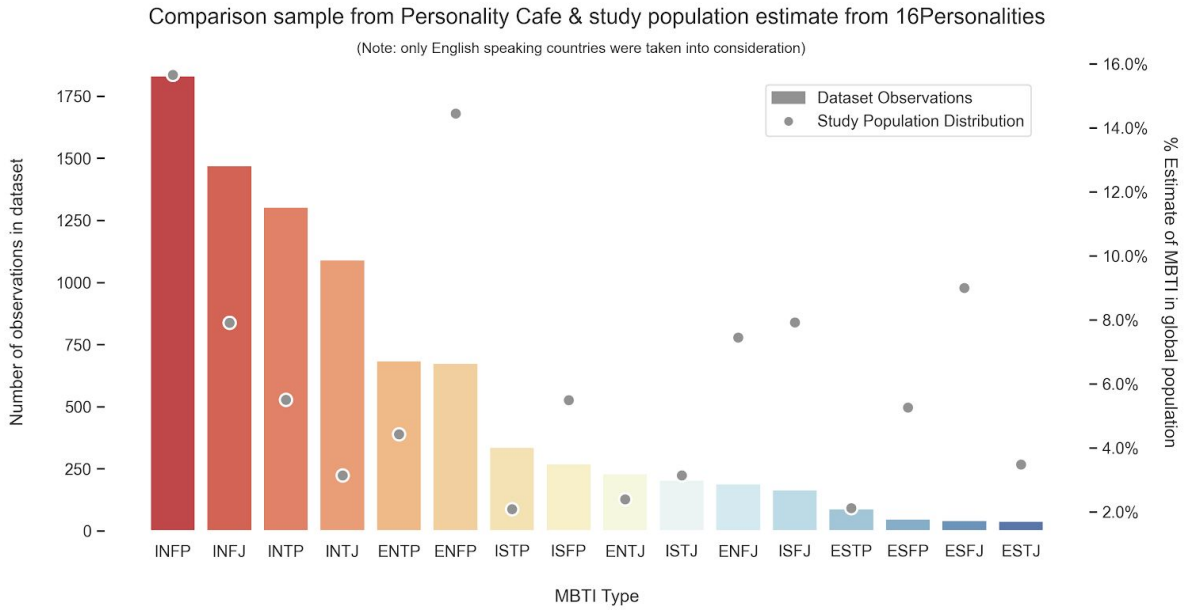
1. The 16 Personality Types (more info)
2. The 4 Keirsey Temperaments (more info)
3. The range of the 4 preferences or "dichotomies"
    a. Extroversion & Introversion (E-I)
    b. Intuition & Sensing*[1] (N-S)
    c. Thinking & Feeling (T-F)
    d. Judging & Perceiving (J-P)

We will use the first level of classification (the 16 types) to describe our data since it is how the dataset came to be and because we have enough observations to make a good analysis. Our goal is to determine if there are relevant differences between the MBTI type in relation to 5 characteristics (number of posts, average number of characters per post, number of hyperlinks, use of emoticons, and self-referencing).

Note: this document shows some figures to explain the main results of the analysis but other interesting insights can be seen in other visualizations uploaded in this GitHub repository

[1] Sensing is sometimes referred to as Observing

SECTION 1 - TARGET VARIABLE



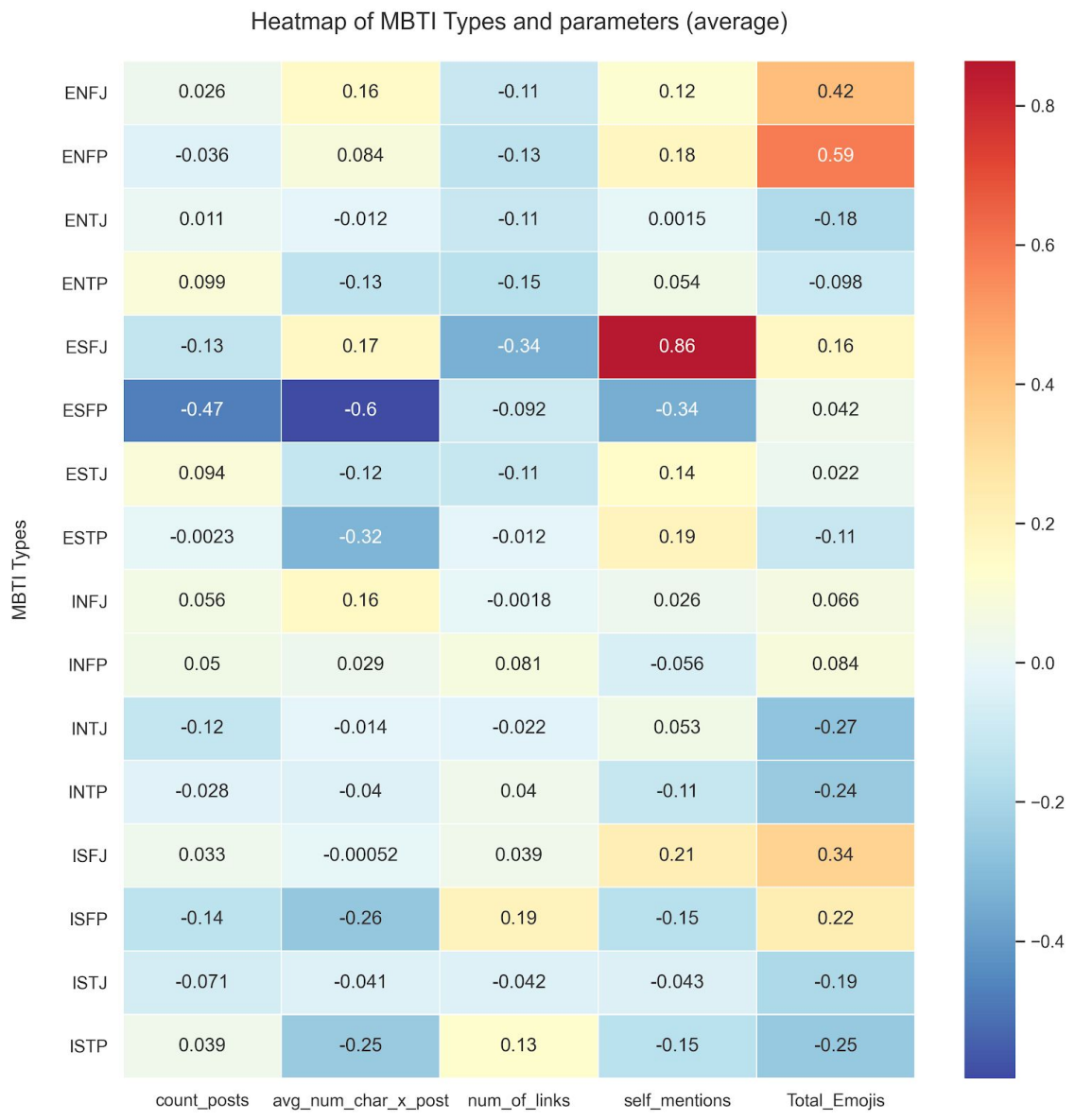
Our target variable is the 16 personality MBTI types. The graph above shows the number of observations this dataset has for each personality type (the bars). Additionally, we have plotted the percentage of the population that falls into each category (dots). The data used to plot these dots comes from the website [16 personalities](). We extracted the values for the biggest English speaking countries (US, UK, Ireland, New Zealand, Australia, and Canada)[2].

There are some interesting differences and information from this first visualization:

- For starters, the distribution of our sample from the blog PersonalityCafe.com differs widely from the sample population we extracted from 16personalities.com. This is especially clear with the ENFP, INTJ, and ESFJ types. We can maybe find some interesting insights looking into the personality types that are over-represented and under-represented in this dataset.
- This graph shows us that intuition types (N) are over-represented in this dataset compared to how we expect the distribution of the general population to behave. The difference is further emphasized when considering Introversion (I).
- The counterpart to this last point comes from sensing types (S) who have the lowest amount of observations. This is further intensified when considering the Extraversion (E) trait.
- Thinking (T) - Feeling (F) and Judging (J) - Perceiving (P) dichotomies do not seem to be strong criteria for the observed differences.

[2] The Personality Cafe forum is mainly English speaking, thus we filtered by English speaking countries. This does not mean that the forum is not used by people from other nationalities.

SECTION 2 - HEATMAP



The heatmap shows the explanatory variables we will use to see if there are significant differences between the 16 MBTI types. These variables are:

- **Count Posts:** This column includes the number of posts a particular person had done when the data was collected.
- **Average Number of Characters per Post:** This information gives us an idea of the average length of the posts someone made
- **Number of Links:** This number is the amount of links someone has shared in their registered posts.
- **Self Mentions:** Refers to the number of times a particular MBTI type refers to itself in its posts. For example: A INTJ saying *"As an INTJ I find this article…"*
- **Total Emojis**: This variable counts the number of emojis a person has used in their posts.

SECTION 3 - STATISTICAL ANALYSIS

We will follow this process to answer our research question. For example:

- **QUESTION:** Do people with different MBTI types write significantly longer or shorter posts?

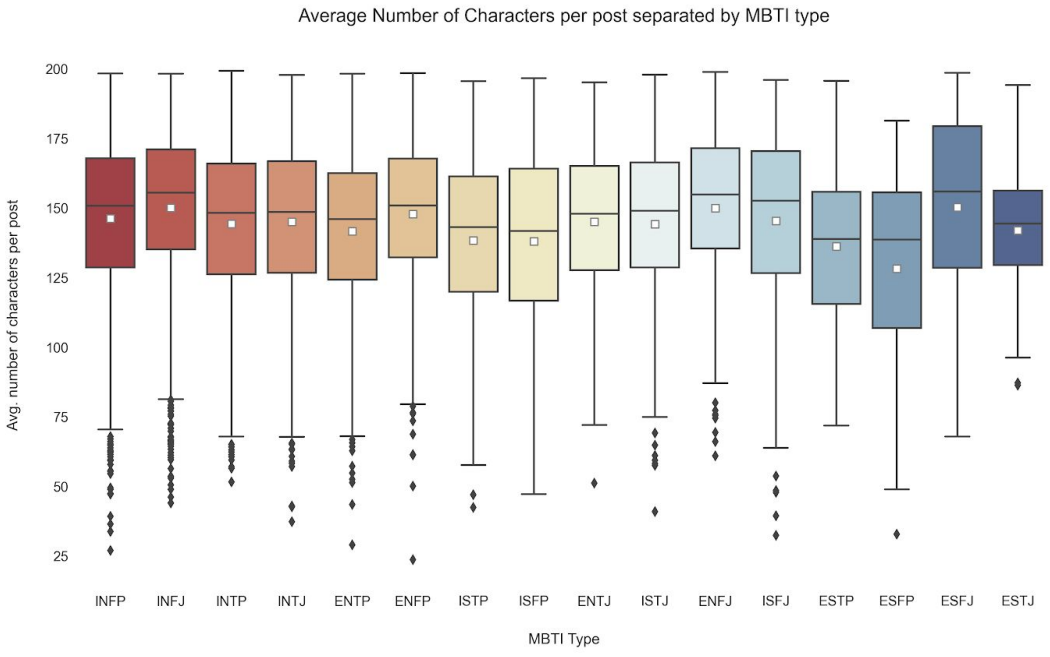
- **PROCESS:** To examine the differences we will check if our data follows a normal distribution, if it does, we will use a one way ANOVA test and if it does not, we will use a Kruskal Wallis test. Since we are doing multiple testing, we can easily end up observing at least one significant result due to chance. If we want to keep our desired significance level, we will need to do some type of correction like Bonferroni or Tukey.

- **ANSWER:** We will end up providing some conclusion to each question.

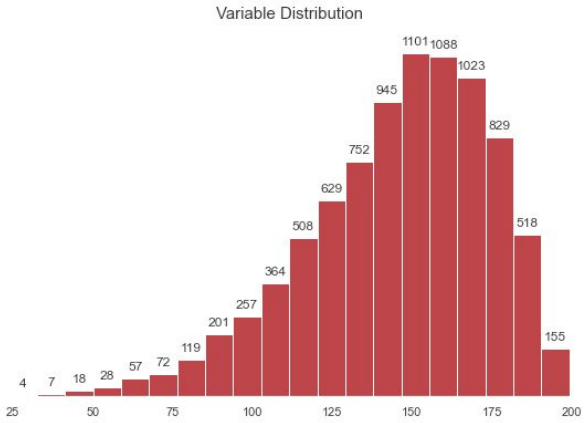
3.1 - Average Number of Characters per post

QUESTION: Do people with different MBTI types write significantly longer or shorter posts?

PROCESS



This boxplot offers some interesting insights, ESFPs seem to show some difference in terms of their mean (white square) being lower than the rest. Additionally, INFJ and ENFJ show higher values in this same statistic.



The distribution seems a little skewed to the left, if we do a Shapiro-Wilk test we get the following statistic and p-value:

- Statistics=0.964, p=0.000 → Sample does not look Gaussian (reject H0)

We will do a Kruskal Wallis nonparametric test to determine if there are differences between the groups. The results of the test are the following:

- H-statistic: 135.6341838645423
- P-Value: 1.6673085613521744e-21
- Reject NULL hypothesis - Significant differences exist between groups.

We need to be careful here and be aware that statistical significance is a function of sample size. Our sample is very large so we can run into statistical significance easily. We can do two things to redress this: do a Dunn's test, and take into account the effect size.

Dunn's Test

Dunn's test is the appropriate nonparametric pairwise multiple comparison procedure when a Kruskal–Wallis test is rejected. You can use the Bonferroni for *post hoc* Dunn's pairwise tests and the results for our sample are the following:

types	ENFJ	ENFP	ENTJ	ENTP	ESFJ	ESFP	ESTJ	ESTP	INFJ	INFP	INTJ	INTP	ISFJ	ISFP	ISTJ	ISTP
ENFJ	-1.0000	1.0000	1.0000	0.0087	1.0000	0.0088	1.0000	0.0030	1.0000	1.0000	1.0000	0.2715	1.0000	0.0020	1.0000	0.0003
ENFP	1.0000	-1.0000	1.0000	0.0175	1.0000	0.0584	1.0000	0.0217	1.0000	1.0000	1.0000	1.0000	1.0000	0.0077	1.0000	0.0004
ENTJ	1.0000	1.0000	-1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.1493	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
ENTP	0.0087	0.0175	1.0000	-1.0000	1.0000	1.0000	1.0000	1.0000	0.0000	0.0127	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
ESFJ	1.0000	1.0000	1.0000	1.0000	-1.0000	0.1993	1.0000	0.3012	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.6861
ESFP	0.0088	0.0584	1.0000	1.0000	0.1993	-1.0000	1.0000	1.0000	0.0018	0.0996	0.3873	0.7272	0.2183	1.0000	1.0000	1.0000
ESTJ	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	-1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
ESTP	0.0030	0.0217	1.0000	1.0000	0.3012	1.0000	1.0000	-1.0000	0.0001	0.0393	0.2797	0.6513	0.2040	1.0000	1.0000	1.0000
INFJ	1.0000	1.0000	0.1493	0.0000	1.0000	0.0018	1.0000	0.0001	-1.0000	0.0032	0.0001	0.0000	1.0000	0.0000	0.4584	0.0000
INFP	1.0000	1.0000	1.0000	0.0127	1.0000	0.0996	1.0000	0.0393	0.0032	-1.0000	1.0000	1.0000	1.0000	0.0106	1.0000	0.0004
INTJ	1.0000	1.0000	1.0000	1.0000	1.0000	0.3873	1.0000	0.2797	0.0001	1.0000	-1.0000	1.0000	1.0000	0.3704	1.0000	0.0439
INTP	0.2715	1.0000	1.0000	1.0000	1.0000	0.7272	1.0000	0.6513	0.0000	1.0000	1.0000	-1.0000	1.0000	1.0000	1.0000	0.2118
ISFJ	1.0000	1.0000	1.0000	1.0000	1.0000	0.2183	1.0000	0.2040	1.0000	1.0000	1.0000	1.0000	-1.0000	0.5842	1.0000	0.1954
ISFP	0.0020	0.0077	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000	0.0106	0.3704	1.0000	0.5842	-1.0000	1.0000	1.0000
ISTJ	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.4584	1.0000	1.0000	1.0000	1.0000	1.0000	-1.0000	1.0000
ISTP	0.0003	0.0004	1.0000	1.0000	0.6861	1.0000	1.0000	1.0000	0.0000	0.0004	0.0439	0.2118	0.1954	1.0000	1.0000	-1.0000

Dunn's test offers pairwise comparisons of our groups. We can see that there is statistical significance in some of the cases. Especifically, INFJ's seem to show more cases of statistical significance in relation to other types, followed by ENFJ, INFP and ISTP

Effect size

It is important to note that statistical significance does not mean practical significance. We can look at the effect size to get more insights on these differences. An effect size refers to the size or magnitude of an effect or result as it would be expected to occur in a population. There are several ways to calculate effect size, a very common one is Hedges' g. The formula for Hedges' g is:

$$\text{Hedges' } g = \frac{M_1 - M_2}{SD^*_{pooled}}$$

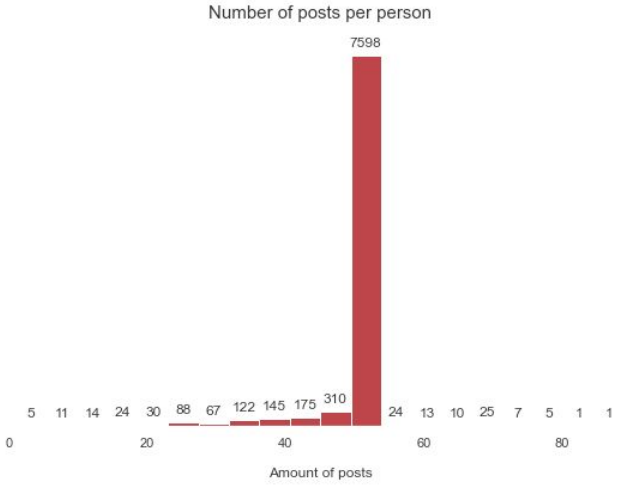
The magnitude of Hedges' g may be interpreted using Cohen's (1988) convention as small > 0.2 (light green color), medium > 0.5 (orange), and large > 0.8 (red).[3] We have compared each mean with the mean of the whole sample. The results are the following:

	Count	Mean	Std	hedges_g
ENFJ	190.000000	149.979981	29.180304	0.160000
ENFP	675.000000	147.890711	26.873478	0.080000
ENTJ	231.000000	145.113648	26.082768	-0.010000
ENTP	685.000000	141.686371	28.684845	-0.130000
ESFJ	42.000000	150.274701	33.550198	0.170000
ESFP	48.000000	128.240393	36.664127	-0.600000
ESTJ	39.000000	142.004465	25.775793	-0.120000
ESTP	89.000000	136.280449	27.408907	-0.320000
INFJ	1470.000000	150.098757	28.240643	0.160000
INFP	1832.000000	146.307324	28.978885	0.030000
INTJ	1091.000000	145.053522	28.346165	-0.010000
INTP	1304.000000	144.299476	27.778352	-0.040000
ISFJ	166.000000	145.449628	33.370462	-0.000000
ISFP	271.000000	138.063837	31.990489	-0.260000
ISTJ	205.000000	144.272117	30.162396	-0.040000
ISTP	337.000000	138.388375	30.033751	-0.250000

[3] Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Erlbaum; Hillsdale, NJ: 1988

ANSWER: The only one that shows a medium effect is the ESFP MBTI type. If we look at Dunn's corrections we can see that the only statistical significant differences are with the ENFJs and INFJs which are the three values that we detected at the beginning.

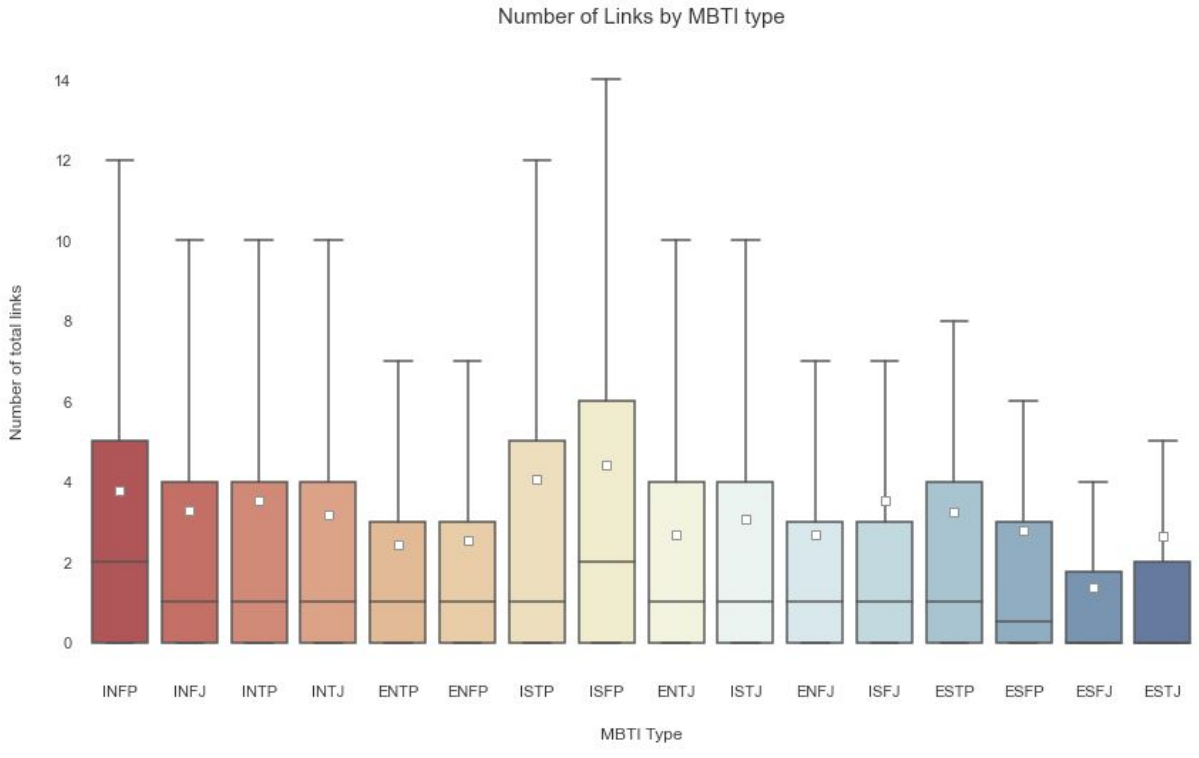
Note: We will follow the same process for each of the other explanatory variables, except for "count of posts", apparently there is no relevant range of values. It is possible that the person who scrapped the Personality Cafe website decided to take 50 posts from each individual directly. We show the histogram here:



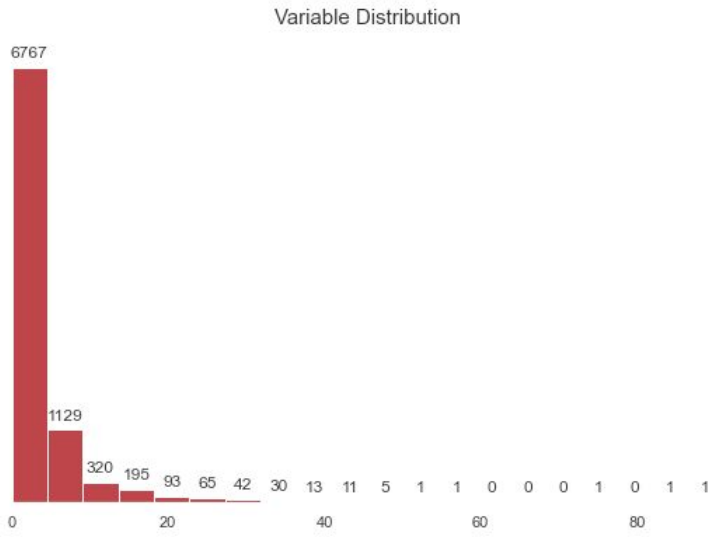
3.2 - Number of Links

QUESTION: Do people with different MBTI types use more or fewer links?

PROCESS



At first glance we see two possible relevant differences, one of them are the ISFPs who seem to use more links than the rest and on the other hand we have the ESFJs who seem to use fewer links. The distribution is the following:



Like with the average number of characters per post we will use a Krustal-Wallis.

- H-statistic: 83.11686593674852
- P-Value: 1.8718416032844887e-11
- Reject NULL hypothesis - Significant differences exist between groups.

Here again we get significant results but let's check Dunn's test:

types	ENFJ	ENFP	ENTJ	ENTP	ESFJ	ESFP	ESTJ	ESTP	INFJ	INFP	INTJ	INTP	ISFJ	ISFP	ISTJ	ISTP
ENFJ	-1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.1312	1.0000	0.6185	1.0000	0.0617	1.0000	0.7553
ENFP	1.0000	-1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.6946	0.0001	1.0000	0.0060	1.0000	0.0029	1.0000	0.0993
ENTJ	1.0000	1.0000	-1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9209	1.0000	1.0000	1.0000	0.3579	1.0000	1.0000
ENTP	1.0000	1.0000	1.0000	-1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0002	1.0000	0.0128	1.0000	0.0049	1.0000	0.1601
ESFJ	1.0000	1.0000	1.0000	1.0000	-1.0000	1.0000	1.0000	0.3732	0.1264	0.0116	0.4406	0.0281	0.9462	0.0040	0.7514	0.0247
ESFP	1.0000	1.0000	1.0000	1.0000	1.0000	-1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
ESTJ	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	-1.0000	1.0000	1.0000	0.9589	1.0000	1.0000	1.0000	0.3550	1.0000	1.0000
ESTP	1.0000	1.0000	1.0000	1.0000	0.3732	1.0000	1.0000	-1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
INFJ	1.0000	0.6946	1.0000	1.0000	0.1264	1.0000	1.0000	1.0000	-1.0000	0.7523	1.0000	1.0000	1.0000	0.9532	1.0000	1.0000
INFP	0.1312	0.0001	0.9209	0.0002	0.0116	1.0000	0.9589	1.0000	0.7523	-1.0000	0.0089	1.0000	1.0000	1.0000	1.0000	1.0000
INTJ	1.0000	1.0000	1.0000	1.0000	0.4406	1.0000	1.0000	1.0000	1.0000	0.0089	-1.0000	0.4172	1.0000	0.0785	1.0000	1.0000
INTP	0.6185	0.0060	1.0000	0.0128	0.0281	1.0000	1.0000	1.0000	1.0000	1.0000	0.4172	-1.0000	1.0000	1.0000	1.0000	1.0000
ISFJ	1.0000	1.0000	1.0000	1.0000	0.9462	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	-1.0000	1.0000	1.0000	1.0000
ISFP	0.0617	0.0029	0.3579	0.0049	0.0040	1.0000	0.3550	1.0000	0.9532	1.0000	0.0785	1.0000	1.0000	-1.0000	1.0000	1.0000
ISTJ	1.0000	1.0000	1.0000	1.0000	0.7514	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	-1.0000	1.0000
ISTP	0.7553	0.0993	1.0000	0.1601	0.0247	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	-1.0000

Here ESFJs are the ones who have more statistical significant differences. However if we look at Hedges g (table below) we see that the difference is quite small (only 0.34).

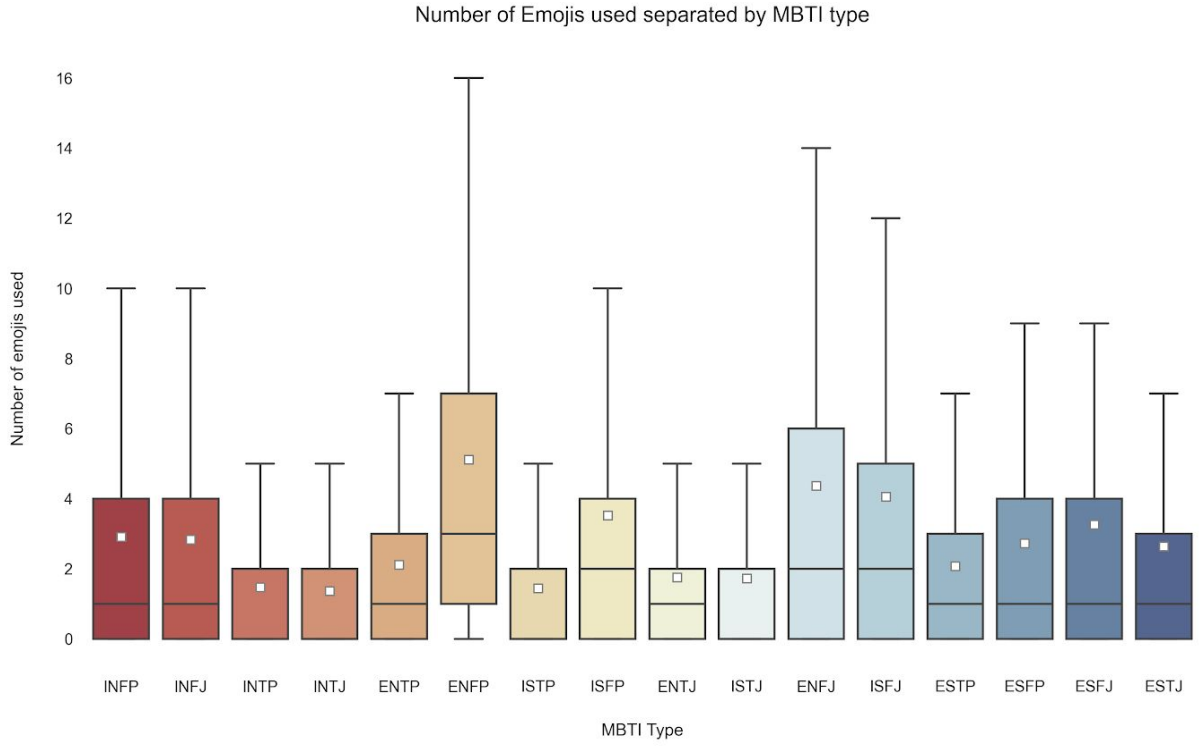
	Count	Mean	Std	hedges_g
ENFJ	190.000000	2.663158	5.062994	-0.110000
ENFP	675.000000	2.522963	4.340745	-0.140000
ENTJ	231.000000	2.658009	4.160202	-0.110000
ENTP	685.000000	2.413139	3.874318	-0.150000
ESFJ	42.000000	1.357143	2.307559	-0.340000
ESFP	48.000000	2.770833	5.413100	-0.090000
ESTJ	39.000000	2.641026	6.466827	-0.110000
ESTP	89.000000	3.235955	6.641752	-0.010000
INFJ	1470.000000	3.293878	5.837426	-0.000000
INFP	1832.000000	3.771288	6.032715	0.080000
INTJ	1091.000000	3.179652	6.073548	-0.020000
INTP	1304.000000	3.538344	5.888280	0.040000
ISFJ	166.000000	3.530120	7.226686	0.040000
ISFP	271.000000	4.416974	7.214142	0.190000
ISTJ	205.000000	3.058537	5.139658	-0.040000
ISTP	337.000000	4.050445	7.559093	0.130000

ANSWER: there does not seem to be a significant effect size in the number of links used by different MBTI types

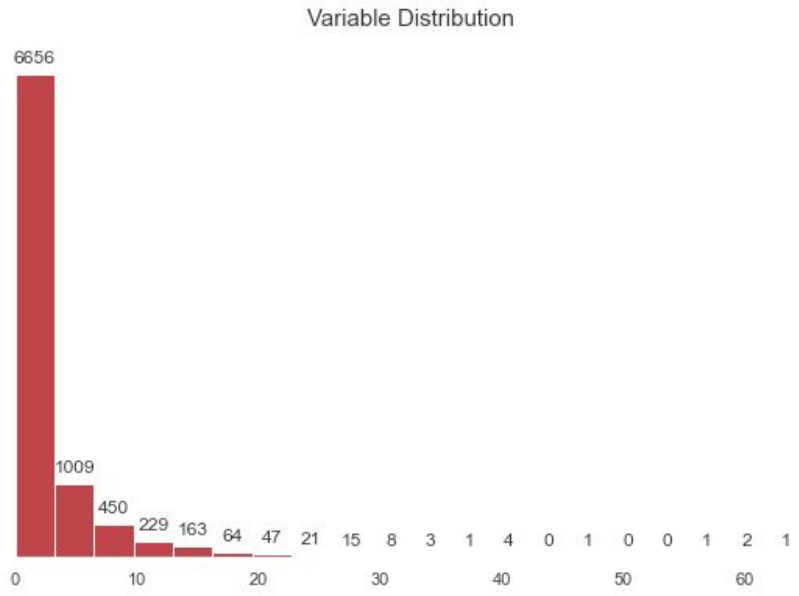
3.3 - Use of Emoticons

QUESTION: Do people with different MBTI types use more or fewer emoticons?

PROCESS



The outlier in this case seems to be ENFP types. Let's see the distribution:



Like in the other cases, the Kruskal Wallis test gives a very low p-value.

- H-statistic: 592.3664819198334
- P-Value: 1.48548010596758e-116
- Reject NULL hypothesis - Significant differences exist between groups.

Dunn's post-hoc pairwise comparison also shows significant differences, especially with the ENFP's.

types	ENFJ	ENFP	ENTJ	ENTP	ESFJ	ESFP	ESTJ	ESTP	INFJ	INFP	INTJ	INTP	ISFJ	ISFP	ISTJ	ISTP
ENFJ	-1.0000	1.0000	0.0000	0.0000	1.0000	1.0000	1.0000	0.0412	0.1419	0.1226	0.0000	0.0000	1.0000	1.0000	0.0000	0.0000
ENFP	1.0000	-1.0000	0.0000	0.0000	0.5829	0.9983	0.9957	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.0612	0.0000	0.0000
ENTJ	0.0000	0.0000	-1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0016	0.0013	1.0000	1.0000	0.0000	0.0003	1.0000	1.0000
ENTP	0.0000	0.0000	1.0000	-1.0000	1.0000	1.0000	1.0000	1.0000	0.0003	0.0001	0.0001	0.0002	0.0000	0.0004	1.0000	0.0692
ESFJ	1.0000	0.5829	1.0000	1.0000	-1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
ESFP	1.0000	0.9983	1.0000	1.0000	1.0000	-1.0000	1.0000	1.0000	1.0000	1.0000	0.3949	0.5571	1.0000	1.0000	1.0000	0.7752
ESTJ	1.0000	0.9957	1.0000	1.0000	1.0000	1.0000	-1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
ESTP	0.0412	0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	-1.0000	1.0000	1.0000	1.0000	1.0000	0.1278	0.9143	1.0000	1.0000
INFJ	0.1419	0.0000	0.0016	0.0003	1.0000	1.0000	1.0000	1.0000	-1.0000	1.0000	0.0000	0.0000	0.8630	1.0000	0.0003	0.0000
INFP	0.1226	0.0000	0.0013	0.0001	1.0000	1.0000	1.0000	1.0000	1.0000	-1.0000	0.0000	0.0000	0.7856	1.0000	0.0002	0.0000
INTJ	0.0000	0.0000	1.0000	0.0001	1.0000	0.3949	1.0000	1.0000	0.0000	0.0000	-1.0000	1.0000	0.0000	0.0000	1.0000	1.0000
INTP	0.0000	0.0000	1.0000	0.0002	1.0000	0.5571	1.0000	1.0000	0.0000	0.0000	1.0000	-1.0000	0.0000	0.0000	1.0000	1.0000
ISFJ	1.0000	1.0000	0.0000	0.0000	1.0000	1.0000	1.0000	0.1278	0.8630	0.7856	0.0000	0.0000	-1.0000	1.0000	0.0000	0.0000
ISFP	1.0000	0.0612	0.0003	0.0004	1.0000	1.0000	1.0000	0.9143	1.0000	1.0000	0.0000	0.0000	1.0000	-1.0000	0.0001	0.0000
ISTJ	0.0000	0.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0003	0.0002	1.0000	1.0000	0.0000	0.0001	-1.0000	1.0000
ISTP	0.0000	0.0000	1.0000	0.0692	1.0000	0.7752	1.0000	1.0000	0.0000	0.0000	1.0000	1.0000	0.0000	0.0000	1.0000	-1.0000

What seems especially interesting is how each type shares similar results with its direct neighbour, for example: ISTx, ISFx, INTx, etc.

The Hedges' g table shows also an interesting case, the ENFP have a medium size effect.

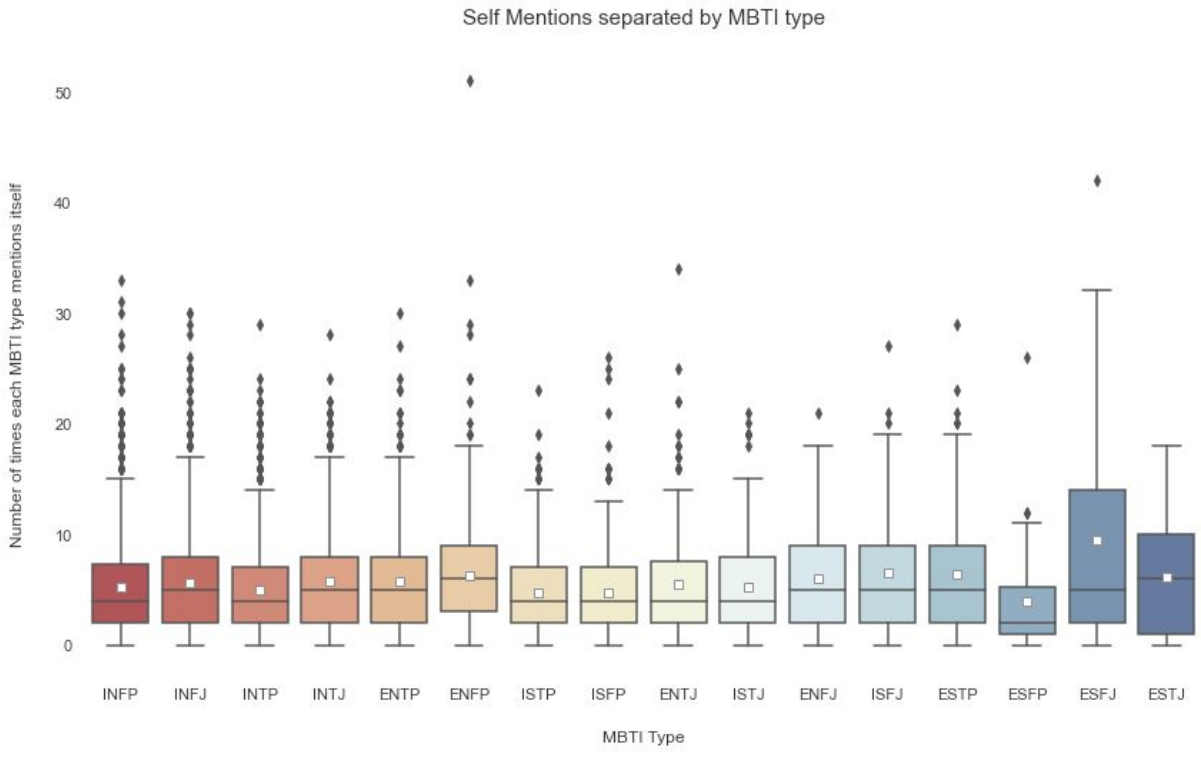
	Count	Mean	Std	hedges_g
ENFJ	190.000000	4.368421	6.075928	0.430000
ENFP	675.000000	5.112593	7.177273	0.600000
ENTJ	231.000000	1.757576	3.015418	-0.180000
ENTP	685.000000	2.115328	3.477958	-0.100000
ESFJ	42.000000	3.261905	5.871884	0.170000
ESFP	48.000000	2.729167	3.609822	0.040000
ESTJ	39.000000	2.641026	3.990291	0.020000
ESTP	89.000000	2.078652	3.419129	-0.110000
INFJ	1470.000000	2.834694	4.172377	0.070000
INFP	1832.000000	2.910480	4.327562	0.090000
INTJ	1091.000000	1.367553	2.568513	-0.280000
INTP	1304.000000	1.470859	3.310479	-0.250000
ISFJ	166.000000	4.054217	5.512588	0.360000
ISFP	271.000000	3.523985	5.906277	0.230000
ISTJ	205.000000	1.726829	3.396341	-0.190000
ISTP	337.000000	1.448071	2.872845	-0.260000

ANSWER: There is a significant difference with a medium effect size in relation to the use of emoticons in the ENFP group.

3.4 Self-Mentions

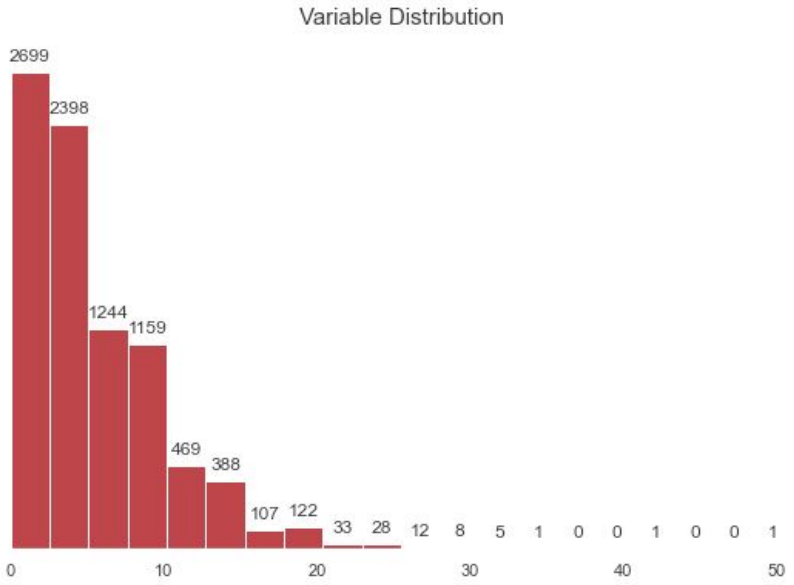
QUESTION: Do people with different MBTI types refer to their own group more or fewer times?

PROCESS



At first glance the ESFJs seem to have higher rates of self-mentioning that the other types, especially in comparison with ESFPs

Let's go deeper by looking at the distribution of the variable:



The Kruskal Wallis test also shows significant results:

- H-statistic: 97.00434932573404
- P-Value: 4.8090704472222525e-14
- Reject NULL hypothesis - Significant differences exist between groups.

Dunn's test shows the ENFP's having a relevant difference with many of the other types. The ESFJs who seemed to have a significant difference in comparison with the others types show no statistical significant difference.

types	ENFJ	ENFP	ENTJ	ENTP	ESFJ	ESFP	ESTJ	ESTP	INFJ	INFP	INTJ	INTP	ISFJ	ISFP	ISTJ	ISTP
ENFJ	-1.0000	1.0000	1.0000	1.0000	1.0000	0.0447	1.0000	1.0000	1.0000	1.0000	1.0000	0.3445	1.0000	0.4768	1.0000	0.8007
ENFP	1.0000	-1.0000	0.1568	1.0000	1.0000	0.0008	1.0000	1.0000	0.0066	0.0000	0.1355	0.0000	1.0000	0.0000	0.0818	0.0000
ENTJ	1.0000	0.1568	-1.0000	1.0000	1.0000	0.8367	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
ENTP	1.0000	1.0000	1.0000	-1.0000	1.0000	0.0400	1.0000	1.0000	1.0000	0.4005	1.0000	0.0055	1.0000	0.1375	1.0000	0.2244
ESFJ	1.0000	1.0000	1.0000	1.0000	-1.0000	0.1368	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
ESFP	0.0447	0.0008	0.8367	0.0400	0.1368	-1.0000	1.0000	0.1515	0.1121	0.6832	0.0596	1.0000	0.0121	1.0000	1.0000	1.0000
ESTJ	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	-1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
ESTP	1.0000	1.0000	1.0000	1.0000	1.0000	0.1515	1.0000	-1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
INFJ	1.0000	0.0066	1.0000	1.0000	1.0000	0.1121	1.0000	1.0000	-1.0000	1.0000	1.0000	0.0226	1.0000	0.6736	1.0000	1.0000
INFP	1.0000	0.0000	1.0000	0.4005	1.0000	0.6832	1.0000	1.0000	1.0000	-1.0000	0.5117	1.0000	0.4871	1.0000	1.0000	1.0000
INTJ	1.0000	0.1355	1.0000	1.0000	1.0000	0.0596	1.0000	1.0000	1.0000	0.5117	-1.0000	0.0040	1.0000	0.2229	1.0000	0.3629
INTP	0.3445	0.0000	1.0000	0.0055	1.0000	1.0000	1.0000	1.0000	0.0226	1.0000	0.0040	-1.0000	0.0436	1.0000	1.0000	1.0000
ISFJ	1.0000	1.0000	1.0000	1.0000	1.0000	0.0121	1.0000	1.0000	1.0000	0.4871	1.0000	0.0436	-1.0000	0.0816	1.0000	0.1372
ISFP	0.4768	0.0000	1.0000	0.1375	1.0000	1.0000	1.0000	1.0000	0.6736	1.0000	0.2229	1.0000	0.0816	-1.0000	1.0000	1.0000
ISTJ	1.0000	0.0818	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	-1.0000	1.0000
ISTP	0.8007	0.0000	1.0000	0.2244	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.3629	1.0000	0.1372	1.0000	1.0000	-1.0000

Interestingly, even if ESFJs do not have an statistically significant difference, they still hold the highest Hedges g.

	Count	Mean	Std	hedges_g
ENFJ	190.000000	6.047368	4.834402	0.120000
ENFP	675.000000	6.315556	4.922178	0.180000
ENTJ	231.000000	5.484848	5.061397	0.000000
ENTP	685.000000	5.729927	4.586084	0.050000
ESFJ	42.000000	9.547619	10.326117	0.870000
ESFP	48.000000	3.875000	4.763686	-0.340000
ESTJ	39.000000	6.153846	5.366269	0.140000
ESTP	89.000000	6.370787	5.736503	0.190000
INFJ	1470.000000	5.597959	4.756620	0.030000
INFP	1832.000000	5.212336	4.525305	-0.060000
INTJ	1091.000000	5.725023	4.709480	0.050000
INTP	1304.000000	4.941718	4.348116	-0.110000
ISFJ	166.000000	6.487952	5.286648	0.220000
ISFP	271.000000	4.752768	4.355272	-0.150000
ISTJ	205.000000	5.273171	4.605777	-0.040000
ISTP	337.000000	4.768546	3.997754	-0.150000

ANSWER: while there are no significant statistical differences, practical significance can be seen in the ESFJ group.

FURTHER ANALYSES

Some topics that could be further analysed are:

a) The use of pronouns by each group (who uses I, me, we more compared to you, they, them?)
b) The use of positive and negative words
c) Most used words with the exception of stop-words