Conclusions based on this sample:

①Survey results vs. MBTI test results

Mind: Tends to be introverted

Energy: Pragmatic

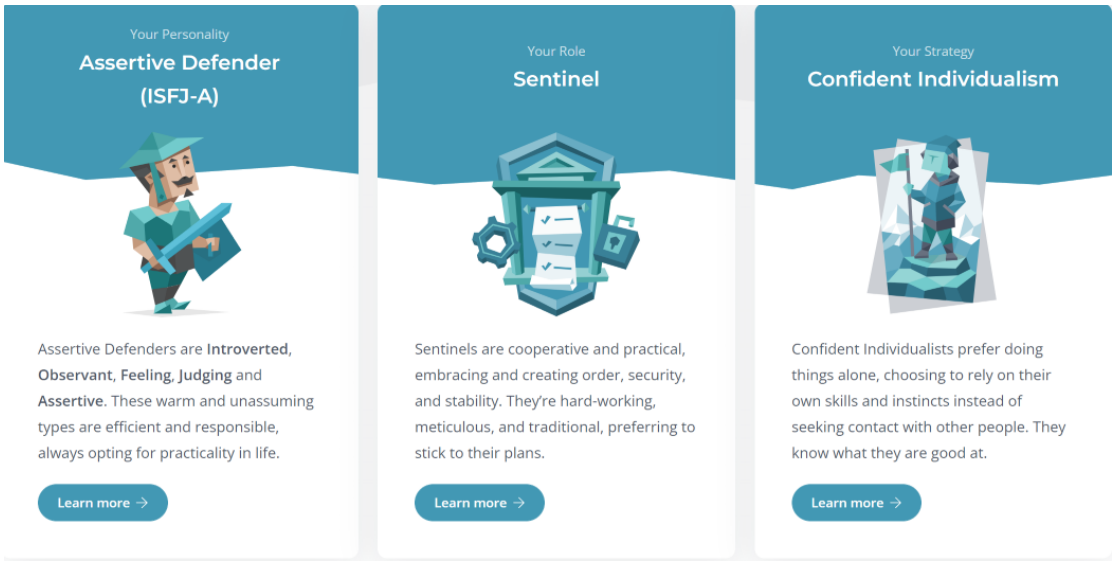
Nature: MBTI test tends to be emotional, survey results tends to be rational

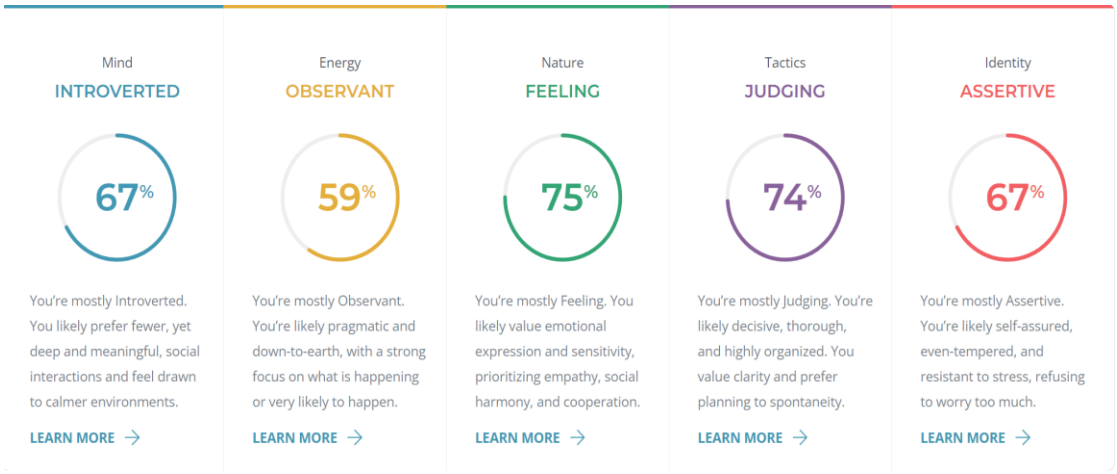
Tactics: Tends to act according to plan

Identity: Tends to be assertive

②in general, there is no difference between men's and women's judgments. For each question, the judgments of men and women are also similar.

MBTI test results:





Personal views:

① For the questionnaire. Each question should be set as a "choice percentage" (e.g., 76%) to increase the accuracy of the results (Prof. Janine Tiefenbruck). When designing the questionnaire, I mainly considered making it as simple as possible, as otherwise few people would be willing to spend a lot of time to complete it. However, Prof. Janine Tiefenbruck suggested that the quality of the questions was more significant. For the next survey, I will focus more on the balance between these two aspects.

② The questions in the questionnaire are actually different from the five items of the MBTI test for comparison. For example, for Energy, Tactics and Identity, I actually only took the most important and simplest characteristics for comparison. For example, in English, there are some differences between "assertive" and "confident". assertive and confident both have the meaning of self-confidence, but assertive specifically refers to Assertive and confident both mean assertive, but assertive means that a person acts confidently, always expresses his or her needs clearly, firmly, and calmly, and makes others take full notice of this characteristic.

③ Different conclusions about Nature. In the MBTI test, it explains "feeling" in this way: "For those with the Feeling trait, decisions tend to be based on the well-being of So, the logic of rational and emotional decisions may just have different starting points, resulting in different solutions, but the goal is the same. For "feeling", it may focus more on the communication aspect; for "thinking", it may focus more on the factual aspect. When dealing with practical problems, it is difficult to say whether the decision is emotional or rational because there are many factors that influence the decision (e.g., people are in different positions or have different positions and take different measures).

④ The problem of sample size. Generally speaking, the larger the sample size and the larger the sample, the closer the experimental results are to the reality. Although there are only 10 samples this time, the results are still good.

Why did you design such an experiment?

I don't think there was any particular reason: it simply seemed fun. At that time, I shared my initial idea with my friends around me, and then a considerable number of them thought the experiment was quite good and fun, so I made a questionnaire. Without their support, I don't know if I would have done such an experiment. Secondly, I also took the opportunity to ask my Stats Prof. Zhang and DSC Prof. Tiefenbruck, which was also a valuable gain in this "journey".

So why? Because Data Science is fun!.... And... STAY HUMBLE!

Idea:

Significance level: 0.05 ($\alpha = 0.05$)

For conclusion 1: Bootstrapping was used to find the confidence interval for the question. If the MBTI test results are in this interval, the null hypothesis cannot be rejected, i.e., the MBTI test results and the survey results support the same conclusion. If it is not, the null hypothesis is rejected.

For conclusion 2: Since the sample size is relatively small, this time a non-parametric test (Nonparametric test) Wilcoxon test was used to compare the differences in judgment between males and females. For each question, males and females are independent samples; for the whole questionnaire, males and females are paired samples.

In the actual code, I also used Student's t-test for your reference. Generally speaking, if the sample size is small, we should check the normality assumption of the data. We can use a histogram or QQ plot to check the distribution of the data, or use the Shapiro-Wilk test (sample size <= 50) and the Kolmogorov-Smirnov test (sample size > 50) to see the size of the p-value. Next, we should check whether the variances of two independent samples are equal. We can use Levene's Test or F test to check this. In general, using two samples t-test mainly tests these two aspects. Since the sample size is relatively small and similar (by looking at the box plot), you can find that neither the use of t-test nor Wilcoxon test will affect the conclusion.

In the report, the first box plot illustrates the data for each question of this sample. The distribution of the Q2 data is more dispersed and the Q11 data is more extreme.

On page six of the report, you can see, in the first line of code, the average score for each question; in the second line of code, the average score is sorted from low to high.

On the bottom part of page seven of the report, you can clearly see the average score per question for males and females.

From page eight to page thirty of the report, the first is the parametric test: t-test, followed by the nonparametric test: Wilcoxon test. the x-axis is gender, and the y-axis is the score for that question. Note that the t-test compares the mean and the Wilcoxon test compares the median.

Please advise me if there are any errors!