

## 1. What is our independent variable? What is our dependent variable?

**Independent variable** (x-axis) would be a tendency of time of how long does it take for participant to finish the task.

**Dependent variable** (y-axis) would be a number of people corresponding to the time they took.

## 2. What is an appropriate set of hypotheses for this task? What kind of statistical test do you expect to perform? Justify your choices.

In human brain, the information process for a reading alphabets will have a priority than that for a colour information. In other words, to process a colour information in human brain will have more steps than reading alphabets, then people will spend more time to say out "the colours of the inks in which the word is printed" on incongruent words than that on congruent words.

Therefore, statistical test will be as below.

**Null Hypothesis ( $H_0$ )** : Time to finish reading congruent words = Time to finish reading incongruent words

**Alternative Hypothesis ( $H_a$ )** : Time to finish reading congruent words < Time to finish reading incongruent words

The two samples have no effect each other, in other words, they are independent samples. For this, negative directional one-tailed t-test having t-critical value 1.711 at  $\alpha = 0.05$  will be performed following central limit theorem, because we do not know the true mean and standard deviation of population; however the sampled mean of congruent words dataset is considered as a population mean.

## 3. Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability.

There are 24 sample data in each test. If  $x$  is a congruent dataset and  $y$  is incongruent dataset, the mean of the congruent dataset,  $\bar{x}$ , is 14.04 with standard deviation for the sample,  $s_x = \frac{\sum x}{n_x - 1} = 3.56$ , and  $\bar{y} = 22.02$  with standard deviation for the sample,  $s_y = \frac{\sum y}{n_y - 1} = 4.80$ .

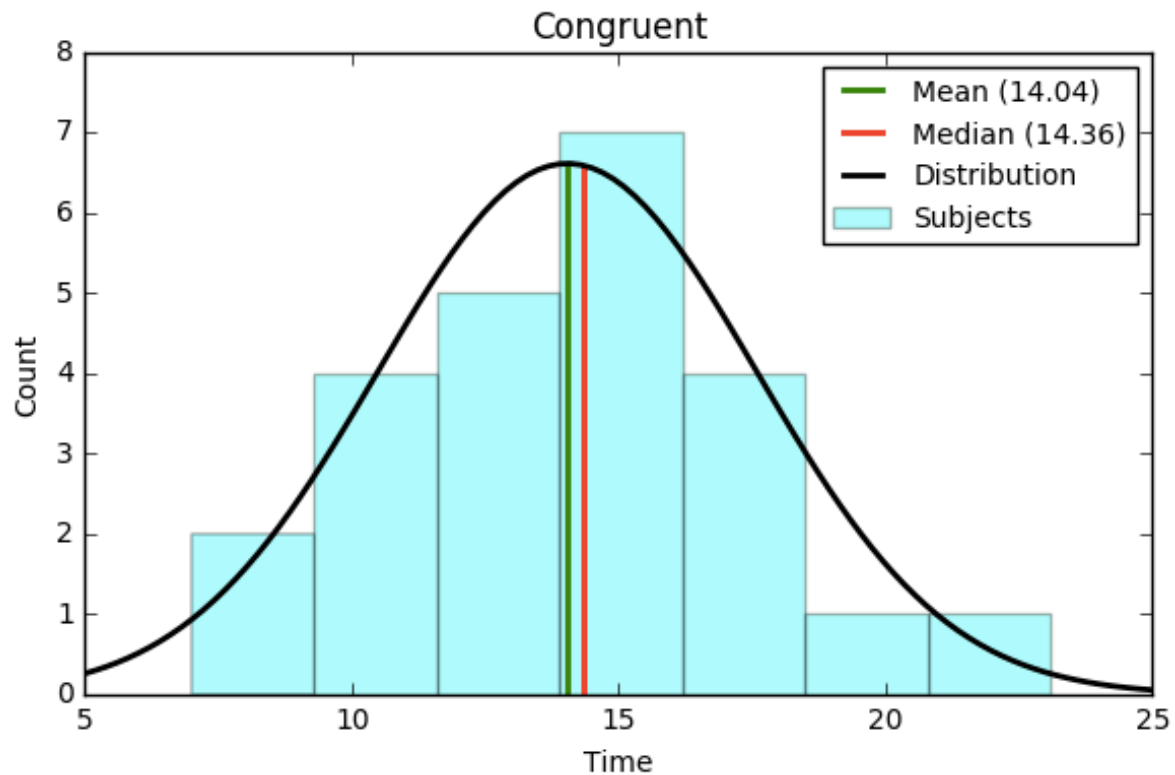
Below is a table of Mean, Standard Deviation, and Median for both sample dataset.

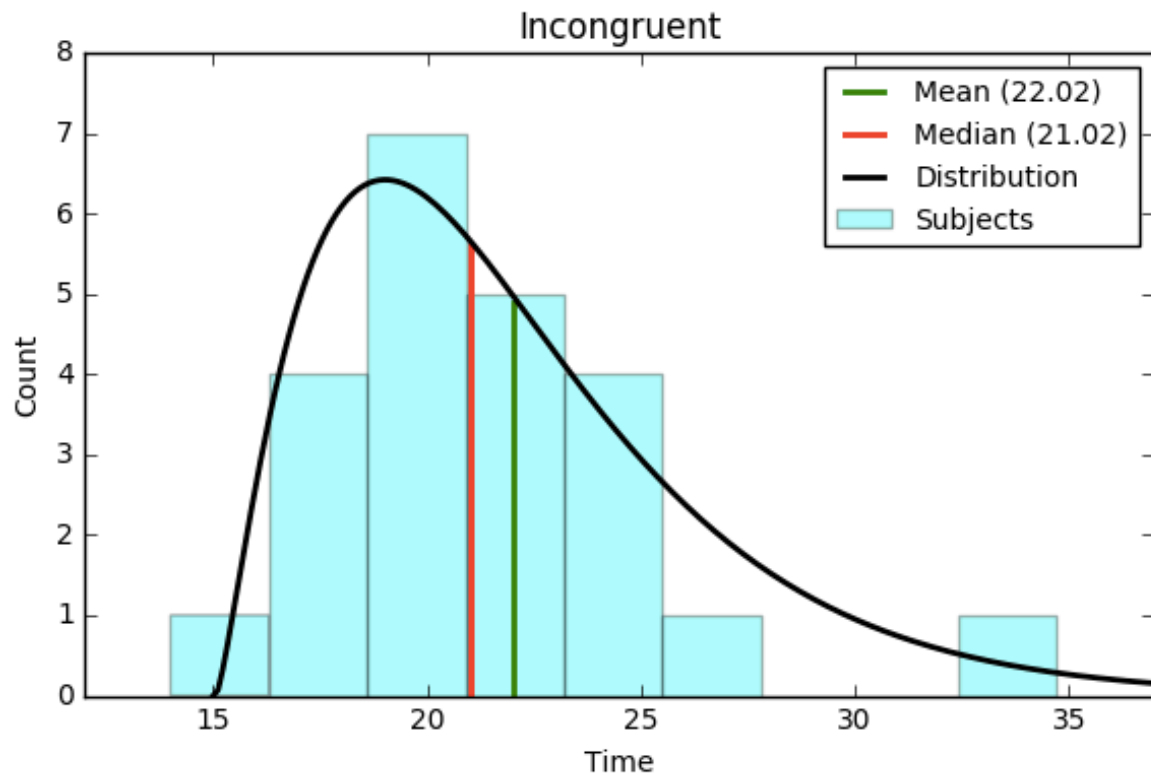
	Congruent	Incongruent	Incongruent - Congruent
Mean	14.04	22.02	8.0
standard deviation	3.56	4.80	in quetion 5
Median	14.36	21.02	7.7

**4. Provide one or two visualizations that show the distribution of the sample data. Write one or two sentences noting what you observe about the plot or plots.**

People tend to take more time to finish the task on the incongruent words than that on the congruent words. In addition, it became wider for the stadard deviation of incongruent tast than that of congruent task. Though the above table tells some descriptive statistics, it is hard to identify wether the hypothesis is significant or not, as well as which kind of central tendency has more appropriate descriptive statistics for each dataset.

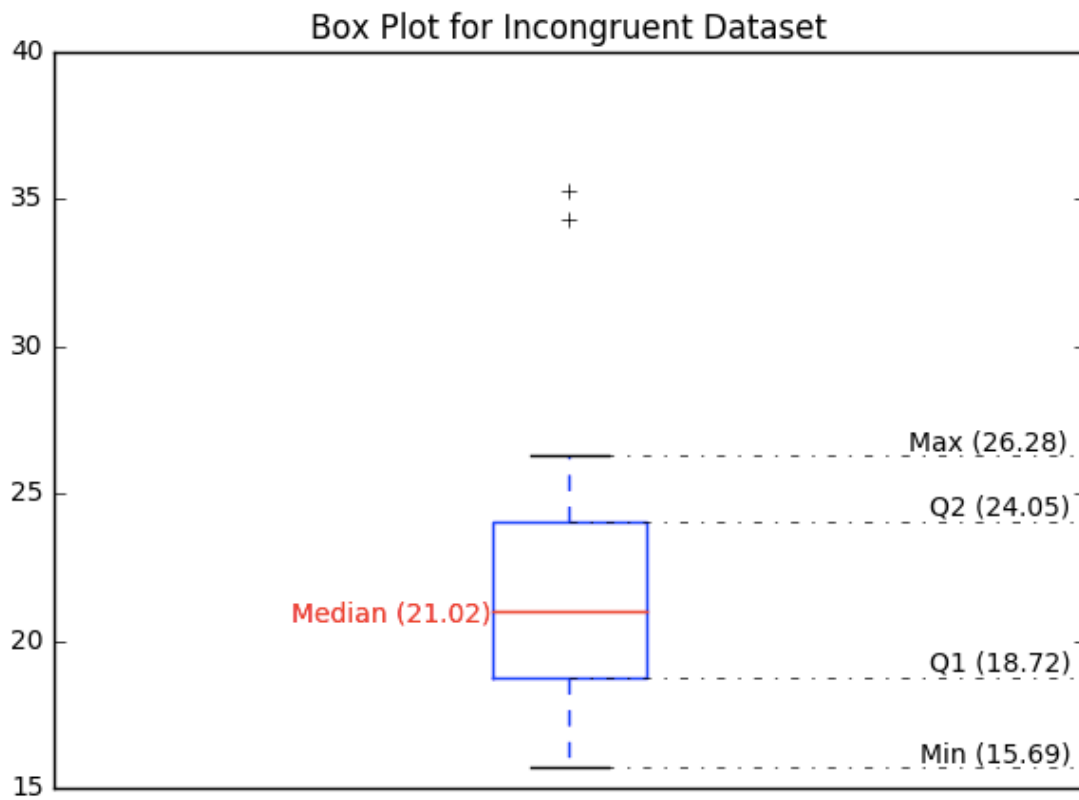
Therefore, it is better to plot a histogram for each dataset so as to find the most descriptive one.





The histogram of the congruent words dataset has 7 bins with 2.3 bin size, and it looks like normal distribution. Theoretically, the mean, median, and mode of its sample dataset should be the same value.

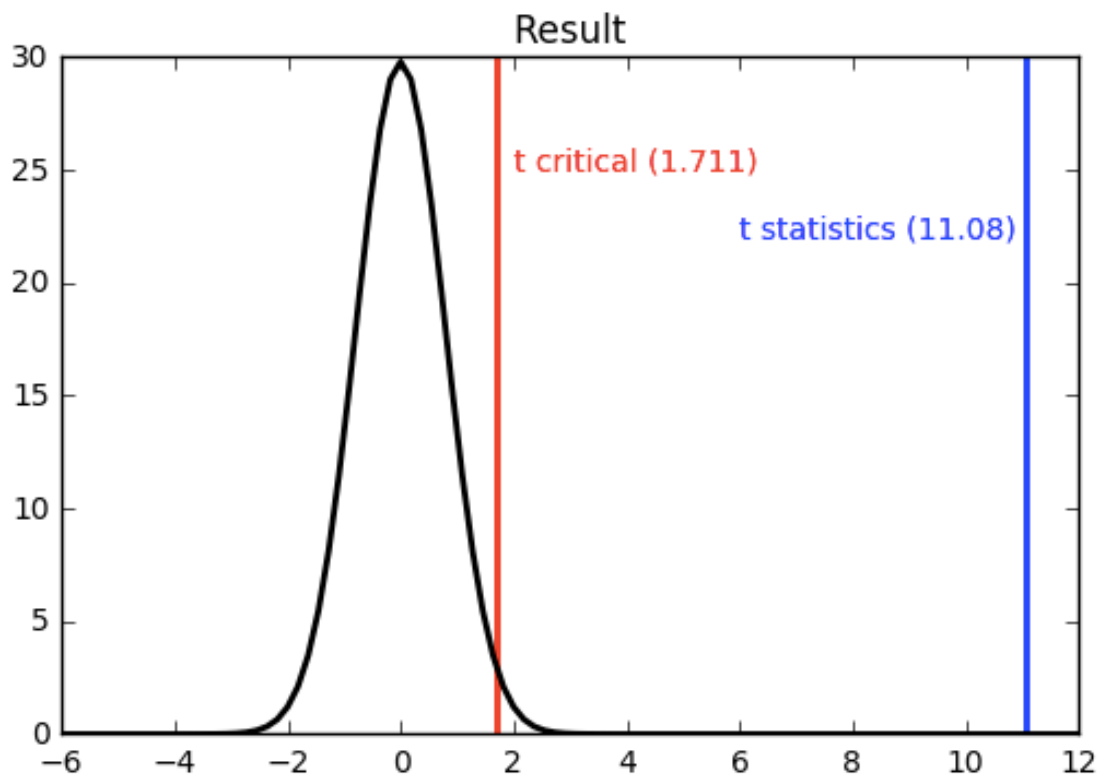
However, though incongruent words dataset has the same bin size, 2.3, as congruent histogram with 9 bins, it seems a positive skewed distribution with one outlier bin and it has an influence the incongruent mean value. The green line (mean value) is slightly on right to the red line (median value). Therefore, IQR must be checked by boxplot to be convinced with the most appropriate kind of central tendency.



Then, it is obvious that the incongruent words dataset has a positive skewed distribution with the 1st quartile (18.72), 2nd quartile (24.05), Min (15.69), Max (26.28), and two outliers, 35.255 and 34.288 respectively.

**5. Now, perform the statistical test and report your results. What is your confidence level and your critical statistic value? Do you reject the null hypothesis or fail to reject it? Come to a conclusion in terms of the experiment task. Did the results match up with your expectations?**

- For this, the sampled mean of congruent words dataset is considered as a population mean ( $\mu = 14.04$ ) with standard deviation ( $s = 3.56$ ).
- Degree of freedom is 23 if  $n = 24$ , and  $t_{critical}$  value is 1.711 if  $\alpha = 0.05$  of one-tailed in plus direction. Standard error of the mean is ( $SEM = \frac{s}{\sqrt{n}} = \frac{3.56}{\sqrt{24}}$ ) is 0.72, which means the sample means to differ from the true population mean by 0.72.
- Mean difference:  $\bar{x} - \mu_{congruent} = 7.98$ .
- t statistics:  $t_{statistics} = \frac{\bar{x} - \mu_{congruent}}{SEM} = \frac{7.98}{0.72} = 11.08$ .
- $t_{statistics}$  falls in the critical region.
- The result is statistically significant.



Therefore, the result says 'reject hypothesis', and which match up with the first assumption, 'the information process for a reading alphabets will have a priority than that for a colour information'.