Questions For Investigation

As a general note, be sure to keep a record of any resources that you use or refer to in the creation of your project. You will need to report your sources as part of the project submission.

1. What is our independent variable? What is our dependent variable?
In this experiment, the independent variables are:
- congruent words condition (colors and words of color match)
- incongruent words condition (colors and words of color do no match)

The dependent variable is the time that participates takes to name the ink colors in equally-sized lists.

2. What is an appropriate set of hypotheses for this task? What kind of statistical test do you expect to perform? Justify your choices.
The appropriate hypothesis could be :
Null hypothesis : the time that takes the congruent test and incongruent test is the same;

$$\mu_{congruent} = \mu_{incongruent}$$

Alternative hypothesis: the time that takes congruent test and incongruent condition is not the same.

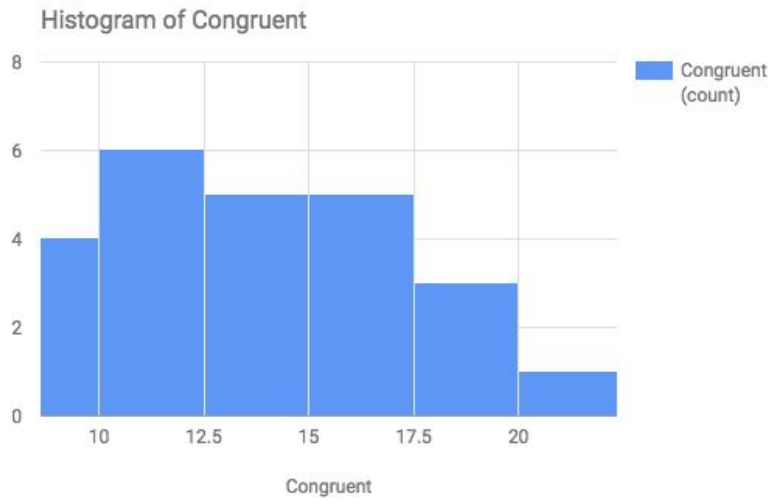
$$\mu_{congruent} \neq \mu_{incongruent}$$

We may perform paired sample student's-T test to see if the two datasets are significantly different from each other. If the calculated t-statistic value is bigger than t-critical value, we reject the null hypothesis; otherwise, we retain the null hypothesis. The reason to perform t-test for this sample is that we have a limited sample size (n<30) and the variables are approximately normally distributed. Meanwhile, as we don't know the population standard deviation, it is best to use t-test for this sample.

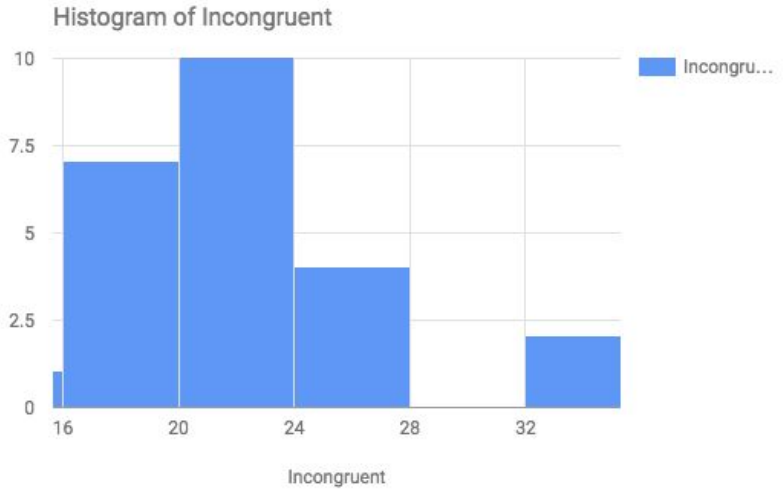
3. Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability.
Descriptive statistics table of the dataset

	Congruent	Incongruent
count	24.000000	24.000000
mean	14.051125	22.015917
std	3.559358	4.797057
min	8.630000	15.687000
25%	11.895250	18.716750
50%	14.356500	21.017500
75%	16.200750	24.051500
max	22.328000	35.255000

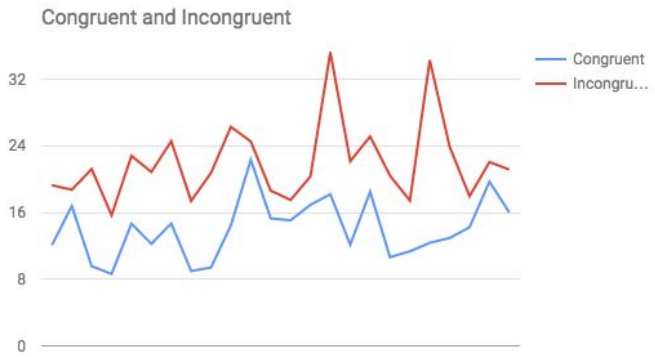
4. Provide one or two visualizations that show the distribution of the sample data. Write one or two sentences noting what you observe about the plot or plots.



Congruent data ranges from 8.63 to 22.328. The median is 14.36.



Incongruent data ranges from 15.687 to 35.255, with most values at the low end. The median is 21.



From the plot of the two datasets, we can see that all data from incongruent condition are higher than data in congruent condition, so we can conclude that the mean and median of incongruent condition is higher than congruent condition.

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?
We conduct a paired t-test:
The mean difference for congruent and incongruent samples.

$$\bar{d} = -7.9647916$$

The standard deviation of the difference

$$s_d = 4.8648269$$

The standard error of the mean difference is

$$SE(\bar{d}) = s_d/\sqrt{n} = 0.993029$$

The t-statistic value is

$$T = \bar{d} / SE(\bar{d}) = -8.0207069$$

The degree of freedom is 23, the two-tailed P value is less than 0.0001, so the difference is considered to be extremely statistically significant. We reject the null hypothesis and conclude that time that takes for congruent test and for incongruent test is different. The results matches my expectations.

6. Optional: What do you think is responsible for the effects observed? Can you think of an alternative or similar task that would result in a similar effect? Some research about the problem will be helpful for thinking about these two questions!
One possible reasons that causes the different is that to recognize the color takes longer time than to recognize the word. Similar task could be to speak out the name of fruits, one test is given in word texts and another in fruit icon pics.