1) **What is our independent variable? What is our dependent variable?**

Independent variable

An independent variable is exactly what it sounds like. It is a variable that stands alone and isn't changed by the other variables you are trying to measure. For example, someone's age might be an independent variable. Other factors (such as what they eat, how much they go to school, how much television they watch) aren't going to change a person's age. In fact, when you are looking for some kind of relationship between variables you are trying to see if the independent variable causes some kind of change in the other variables, or dependent variables.

Dependent variable

A dependent variable is exactly what it sounds like. It is something that depends on other factors. For example, a test score could be a dependent variable because it could change depending on several factors such as how much you studied, how much sleep you got the night before you took the test, or even how hungry you were when you took it. Usually when you are looking for a relationship between two things you are trying to find out what makes the dependent variable change the way it does.

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

Null Hypothesis

The null hypotheses implies that the test is one-tailed since the mean color recognition time for congruent words is equal to or more than that of incongruent words.

$H_0: \mu_C \geq \mu_I$

where μ is a population mean, the subscript "C" represents the congruent words condition, and the subscript "I" represents the incongruent words condition.

Alternative hypothesis

The alternative hypothesis, denoted by H1 or Ha, is the hypothesis that sample observations are influenced by some non-random cause. It should state that the mean of congruent words is less than that of the mean of incongruent words.

$H_A: \mu_C < \mu_I$

One Tailed Dependent Sample t-test

A one-tailed, dependent samples t-test comparing the difference in means (the time to name the ink colors for congruent words and incongruent words) should be performed. With this test, we seek to determine whether there is enough evidence in the provided sample of data to infer that the congruent words mean color recognition time is less than the incongruent words mean color recognition time for the entire population and not just the sample data.

Example

A t-test is appropriate because the population variance is unknown and the sample size is less than 30. When the sample size is less than 30, the sample data no longer approximate a normal

distribution, which makes the use of a Z-value inappropriate.[1] The following assumptions are required for t-tests for dependent means

Characteristics:

Interval or ratio scale of measurement (approximately interval)

Random sampling from a defined population

- Samples or sets of data used to produce the difference scores are linked in the population through repeated measurement, natural association, or matching
- Scores are normally distributed in the population; difference scores are normally distributed

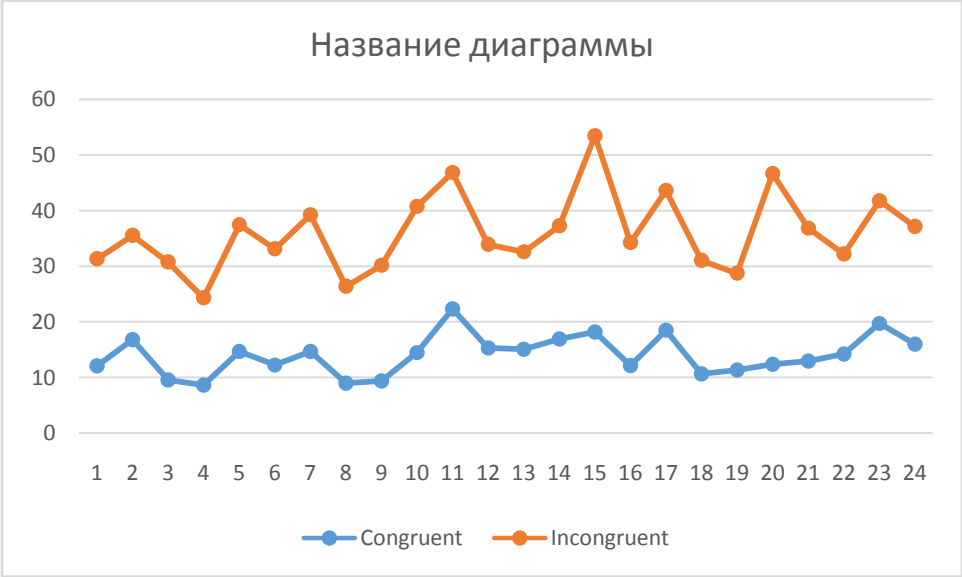
A one-tailed test is appropriate under the assumption that incongruent word conditions will not improve recognition times, which is intuitive. The one-tailed test allows for a more scrutinous examination of the negative impact of incongruent word conditions on recognition times.

The t-test should be of the dependent samples variety because the same subject is exposed to two conditions and tested for each, which are the defining criteria for "within-subjects" or "repeated-measures" statistical tests.

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

Statistic	Congruent	Incongruent	Difference
n	24	24	24
$\bar{X}$	14.05	22.02	-7.96
Median	14.36	21.02	-7.67
SD^2	12.67	23.01	23.67
SD	3.56	4.80	4.86
SE	0.73	0.98	0.99

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.



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?

One Tailed t-test for 99% (α)= 0.01

Degree Of Freedom = 23

Critical statistics(t_{crit})= -2.50

T - statistics (t) = -8.02

P-value = < .0001

References:

1) https://classroom.udacity.com/nanodegrees/nd009-in-basic/parts/2d574aac-201b-4501-96b6-03107ef64d89
2) http://www.rpubs.com/zxia924/stroopeffect
3) https://rpubs.com/georgeliu/nanop1