UDACITY PROJECT

SUBMITTED BY: Akanksha Goel

DATE: 16 May,2017

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

**Dependent Variable:**  The time it takes to name the ink colors in equally-sized lists.

**Independent Variable** : The two independent conditions that is chosen whether the word and color is congruent or incongruent.

**Source**: Given.

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

Ho (null hypothesis) : that there will be no affect in the time taken while reading incongruent words .

**µ(congruent) ≥µ(incongruent)**

**(population mean of time taken will remain same to population mean when used congruent or incongruent set.)**

Ha (alternative hypothesis): that there will be increase in time taken while reading incongruent words i.e

**µ(congruent) <µ(incongruent)**

**(population mean will increase as you use the incongruent set instead of congruent set)**

In this **t test is used** to analyse hypothesis as we **don’t have population standard deviation σ** and don’t have **sample size greater than 30**.

We assume that the distributions are Gaussian.

This test is applied on these **dependent samples** because same population is used to read congruent set and incongruent set. **Time is measured in two conditions on the same population .**

This hypothesis testing will be conducted using **one-tail test**.

This statistical analysis is being done because time used in reading incongruent words is greater then time required to read congruent words as per the stroop’s effect.

As **correlation cannot be considered as causation of effect** therefore inferential statistics is being used to infer results using experiment.

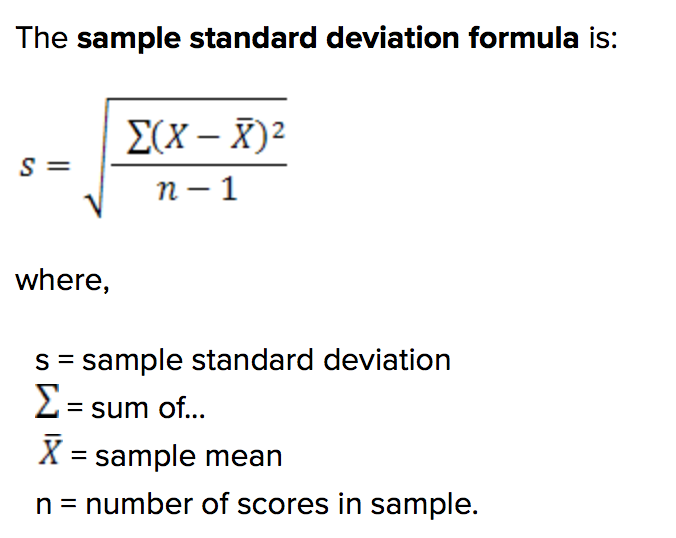
**Source**: https://en.wikipedia.org/wiki/Stroop\_effect

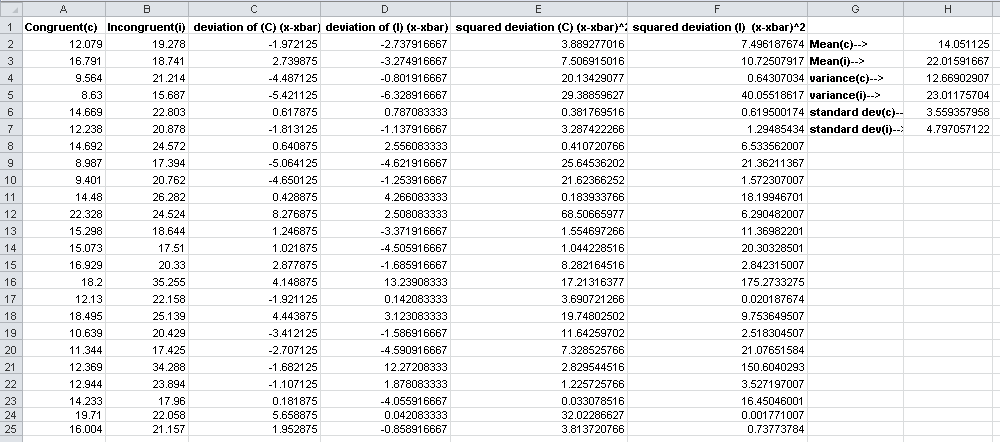
1. **Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability.**

**Measure of central tendency :** Mean(congruent) = Average time taken to read congruent words = .

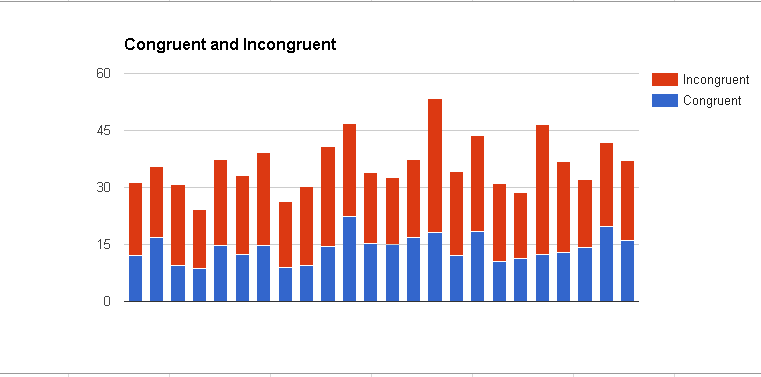
Mean(incongruent) = Average time taken to read incongruent words = Related image

**Measure of variability :**



.

1. **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.**



X Axis represent users.

Y Axis represents time taken by respective user.

For sample:

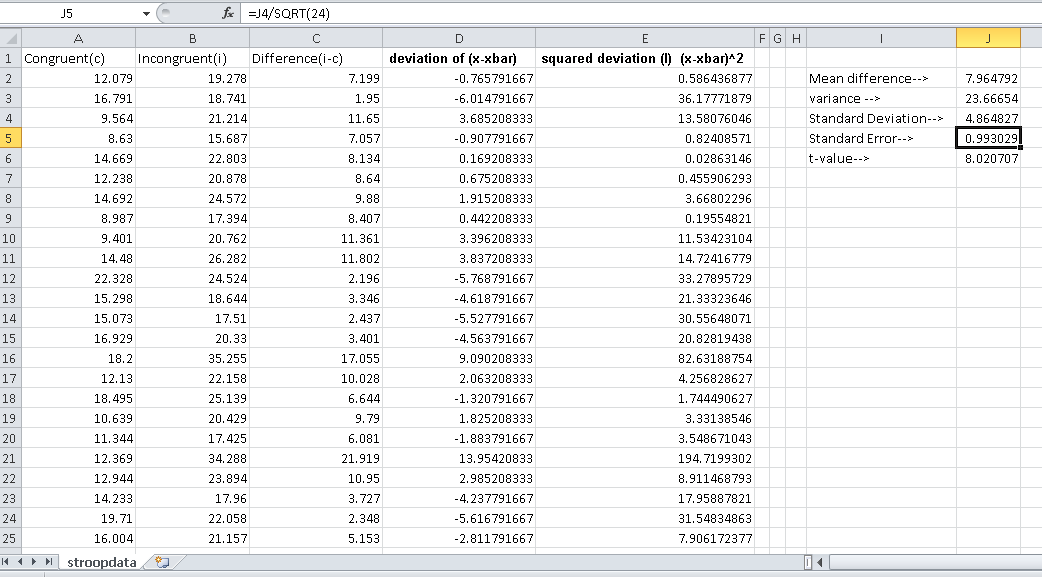
Time taken in reading incongruent words > Time taken while reading congruent words.

For population:

The chart shows that time taken while reading incongruent words may be more than that of reading congruent words for user.

1. **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?**

* Using t-distribution we check that the independent variable will have an effect on dependent variable using 1 tail test.
* Let the α value for distribution be 0.05.
* n=sample size=24
* **degree of freedom**: 23
* Finding the **t-critical value**:1.714 (using t-table)
* Mean Difference : 7.964
* Standard Deviation of difference :4.86



* standard Error = = = 0.9930
* t statistical value = = = 8.020
* P-value for this t value must be less then .05 because t value is greater then t-critical value.

P = 0.05

P<0.05

1.714 8.02

* As t (statistical value ) is greater than t-critical value therefore

The t value lies in critical region therefore we will reject null hypothesis and accept alternative Hypothesis .

**µ(congruent) <µ(incongruent) (population mean)**

Hence it is evident that bringing incongruent words increase the time taken by user.

**Yes, the result came out to be totally different to what we expected. *The time taken while reading incongruent words is more than the time required to read congruent words .***

* For finding confidence interval let α be 0.05.

Now t-critical value be ± 2.069 .(using t-table)

Standard error : .9930

Mean difference : 22.01 – 14.05 = 7.96 .

**95% confidence interval : mean difference ± (t-critical \* Standard Error) = ( 5.91, 10.008).**

1. **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!**

There are two theories that may explain the Stroop effect:

1. Speed of Processing Theory: the interference occurs because words are read faster than colors are named.
2. Selective Attention Theory: the interference occurs because naming colors requires more attention than reading words.

**Alternative Task :** Colouring the words half with one colour and other half with different colour and reading the word instead of colour