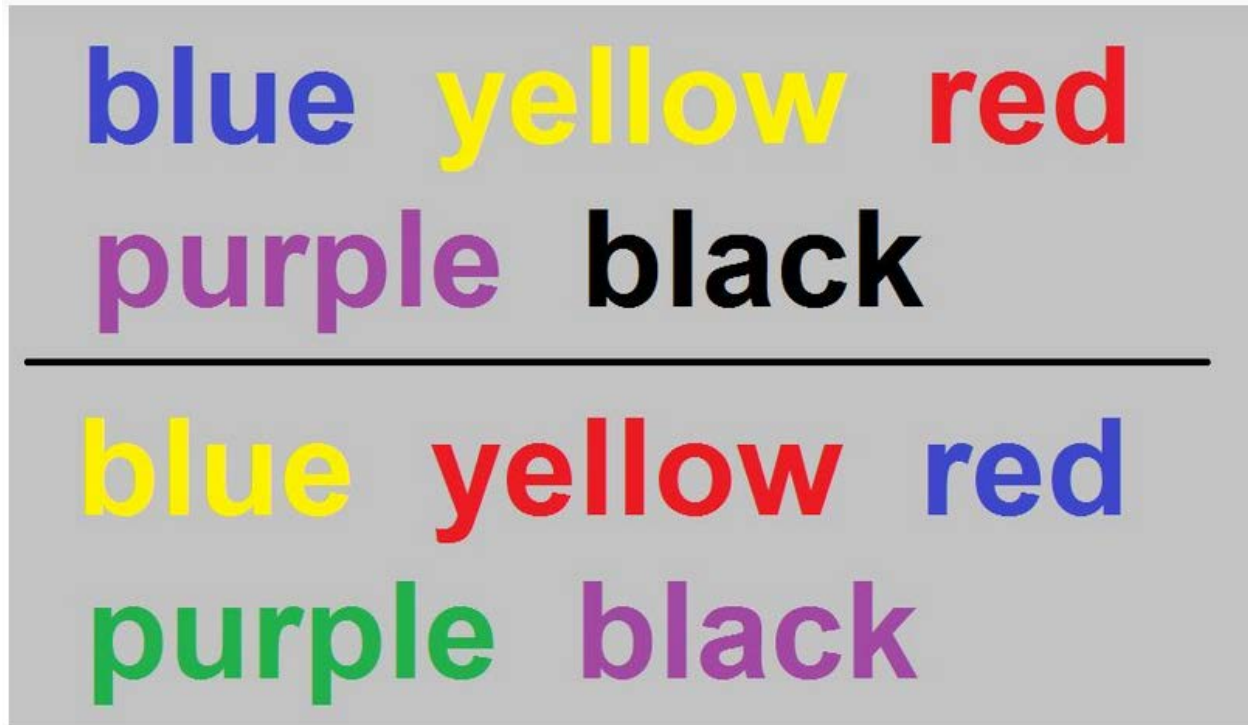


Statistics: The Science of Decisions



In this project, we investigate the psychological phenomenon known as Stroop effect. The **Stroop effect** is a demonstration of interference in the reaction time of a task. When the name of a color (e.g., "blue", "green", or "red") is printed in a color that is not denoted by the name (e.g., the word "red" printed in blue ink instead of red ink), naming the color of the word takes longer and is more prone to errors than when the color of the ink matches the name of the color.

Below are the investigations performed and observation of the sample dataset which contains data from the experiment performed in 2 conditions - congruent and incongruent.

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

Independent variable is the word color congruency that is being manipulated in the experiment.

Dependent variable is the reaction time of saying the words out loud.

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

Let the population mean of reaction time for the congruent words be μ_c and the population mean reaction time for incongruent words be μ_i .

The null hypothesis is the time to name the ink colors of incongruent words is less than or equal to the time to name the ink colors congruent words.

The alternate hypothesis is the time to name the ink colors of incongruent words is greater than the time to name the ink colors congruent words.

$$H_0: \mu_i - \mu_c \leq 0$$

$$H_1: \mu_i - \mu_c > 0$$

The variances and standard deviation of the sample data are not known so we will perform a T-test. We will perform a paired sample test as the same group of participants performed the task for both congruent and incongruent words.

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

My Results:

Congruent -16.145

Incongruent - 21.469

Below are some descriptive statistics of the dataset:

	Congruent	Incongruent
Count	24	24
Min	8.63	15.687
Max	22.328	35.255
Mean	14.051	22.016
Standard Deviation	3.559358	4.797057
1 st Quartile	11.895250	18.716750
2 nd Quartile	14.356500	21.017500
3 rd Quartile	16.200750	24.051500

In the above dataset **measure of central tendency** is the mean. For congruent the value is 14.051 and for incongruent is 22.016.

Few **measures of variability** are below:

	Congruent	Incongruent
Range	13.698	19.568
Standard Deviation	3.559358	4.797057

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.

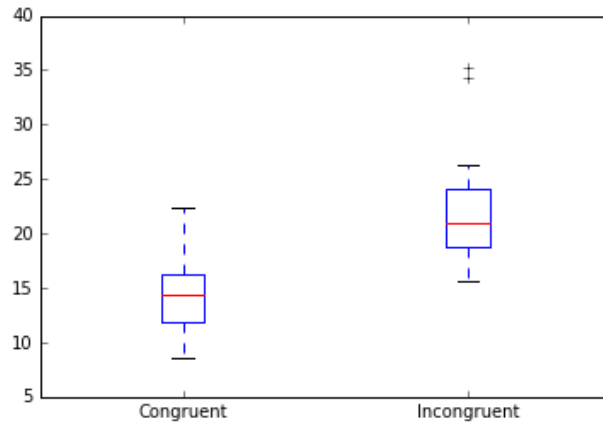


Fig 1: Box-Plot

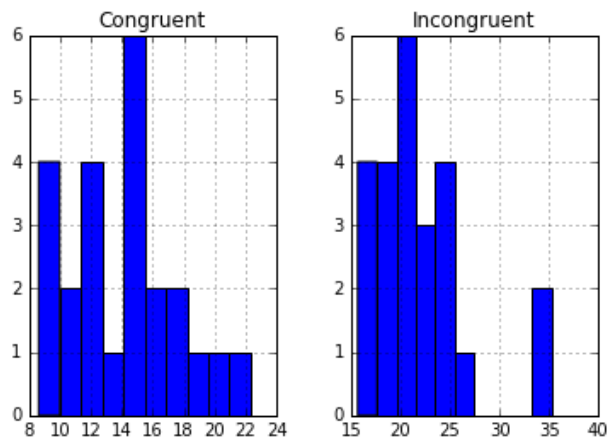


Fig 2: Histogram

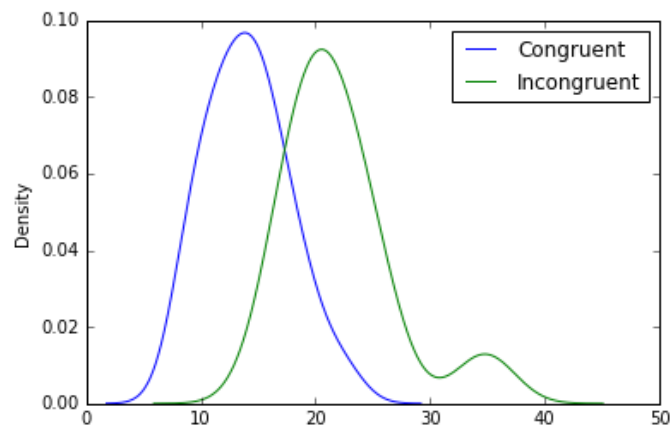


Fig 3: Density Plot

From the box plot we can observe that there are few outliers present in incongruent data. Also, there is difference between the median value of congruent and incongruent observations.

From the histogram and density plot, we can observe there is small skewness in the distribution of data for incongruent observations but the mean is close to the peak of the distribution. The congruent observations are normally distributed.

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?

Below are the findings of the paired sample t-test:

T-statistic = 8.0207

p-value = 4.1030005857111781e-08

since p-value is less than 0.05, the result is significant. Hence, we reject the null hypothesis.

Conclusion:

Based on the hypothesis testing (paired t-test) performed above we conclude that naming of color of words takes longer for incongruent words as compared to congruent words.

References:

<https://pandas.pydata.org/pandas-docs/stable/visualization.html>

https://en.wikipedia.org/wiki/Stroop_effect

Business Analytics by James R. Evans

https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.ttest_rel.html