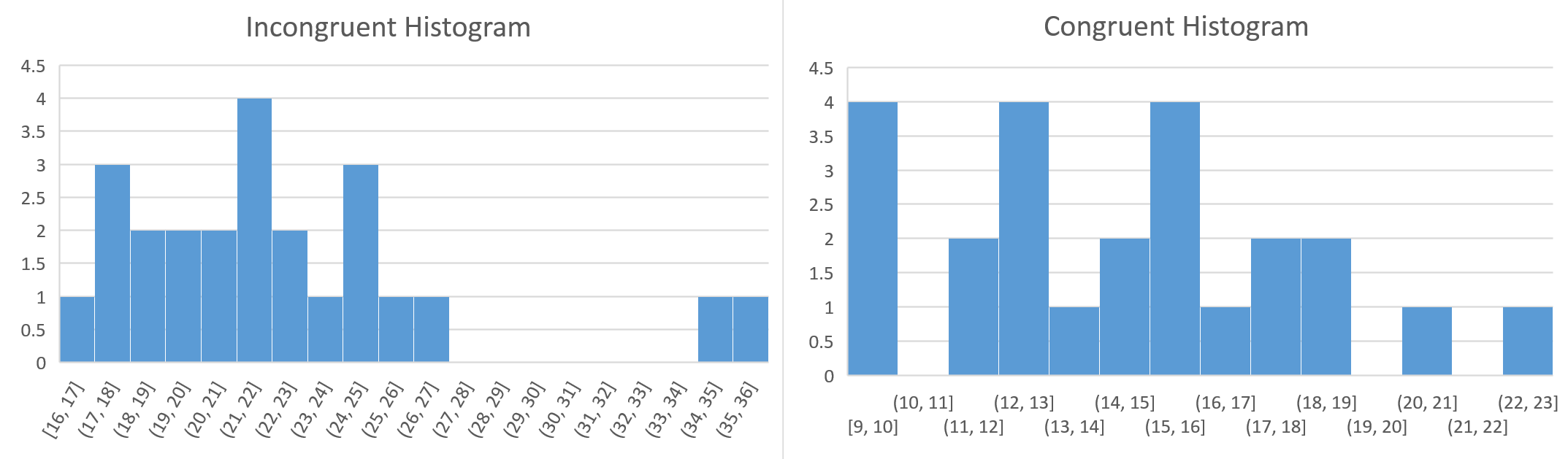
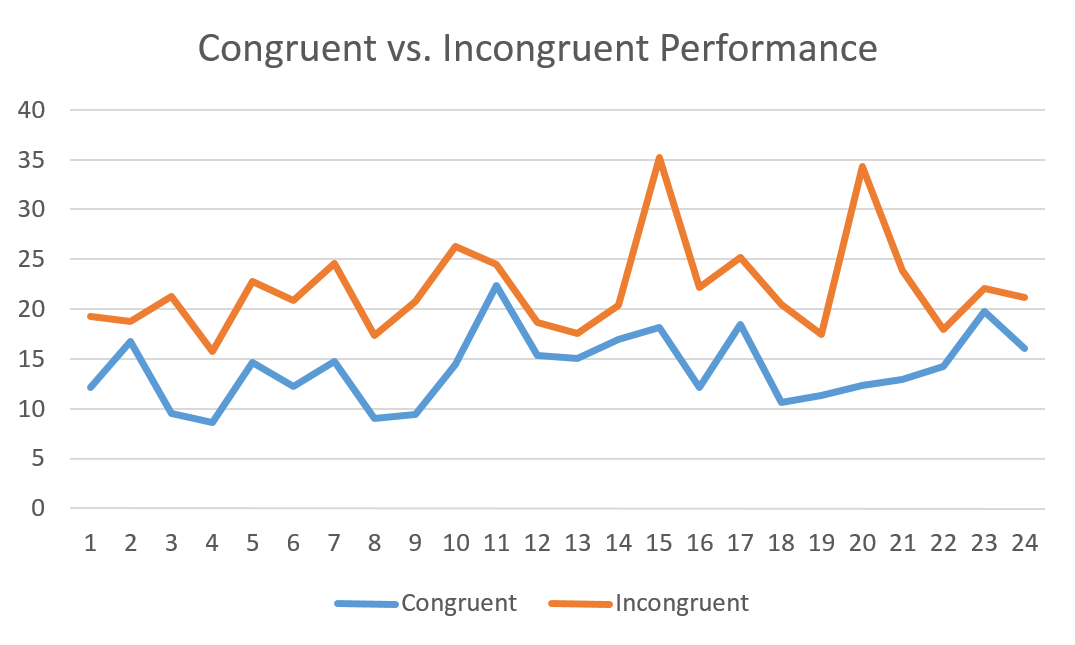
**Statistics Concepts**

* Lesson 1
  + Constructs and operational definitions: Start at [the beginning of Lesson 1](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4601188734/m-4625128561) and continue until you reach the video "Data".
  + Interpreting scatter plots: Start at the video ["Data"](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4601188734/m-56181653) and continue until you reach the video "Benefits of Surveys".
  + How to make a scatter plot in Google spreadsheets: Watch the video ["Katie's Hand"](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4601188734/e-56240821/m-56181694).
    - See also [these instructions](https://support.google.com/docs/answer/140940?hl=en&ref_topic=30238).
  + Correlation vs. causation: Start at the video ["Visualize Relationship"](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4601188734/e-210129179/m-210129180) and continue until you reach the video "Benefits of Surveys".
  + Different types of studies: Start at the video ["Causal Inference"](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4601188734/m-209236000) and continue to the end of Lesson 1.
  + Terminology (population, sample, parameter, and statistic): Watch the video ["Same Scores" (solution)](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4601188734/e-56533421/m-56181661).
* Lesson 2
  + Absolute and relative frequency, proportions, and percentages: Start at [the beginning of Lesson 2](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4574374276/e-59610756/m-233118803)and continue until you reach the video "Number of Rows".
    - See also [this summary](http://www.abs.gov.au/websitedbs/a3121120.nsf/home/statistical+language+-+describing+frequencies).
  + Histograms and binning: Start at the video ["Number of Rows"](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4574374276/e-59706478/m-59645470) and continue to the end of Lesson 2.
* Google Spreadsheet Tutorial Lesson
  + How to perform calculations in Google spreadsheets: Watch [the entire lesson](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4595028745/m-116474982).
    - See also [this tutorial](http://www.gcflearnfree.org/googlespreadsheets/12).
* Lesson 3
  + Mode: Start at [the beginning of Lesson 3](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4615238546/e-77367476/m-77345002) and continue until you reach the video "Find the Mean".
  + Mean: Start at the video ["Find the Mean"](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4615238546/e-77314347/m-77345030) and continue until you reach the video "UNC".
  + Median: Start at the video ["UNC"](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4615238546/m-77345048) and continue to the end of Lesson 3.
  + Different shapes of distributions - normal, uniform, bimodal, and skewed: Covered throughout[Lesson 3](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4615238546/e-77367476/m-77345002).
    - See also [this page](http://www.mathbootcamps.com/common-shapes-of-distributions/). (A normal distribution is a type of bell-shaped distribution.)
* Lesson 4
  + Range: Start at [the beginning of Lesson 4](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4612348621/e-83877448/m-83664002) and continue until you reach the video "Mark Z the Outlier".
  + Inter-Quartile Range and Boxplots: Start at the video ["Mark Z the Outlier"](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4612348621/e-83877443/m-83664016) and continue until you reach the video "Problem with IQR".
  + Variance and standard deviation: Start at the video ["Problem with IQR"](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4612348621/m-83664040) and continue to the end of Lesson 4.
  + Bessel's correction: Start at the video ["Sample SD"](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4612348621/e-83934429/m-84715735) and continue to the end of Lesson 4.
    - See also [this article](https://en.wikipedia.org/wiki/Bessel%27s_correction).
  + Normal distribution 68-95 rule: Watch the video ["Point of SD"](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4612348621/m-83664096).
* Lesson 5
  + z-scores: Watch [the entire lesson](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4575424090/e-92197576/m-92314567).
* Lesson 6
  + Using the z-score table: Watch [the entire lesson](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4588768731/m-102088771).
* Lesson 7
  + Sampling distributions and the Central Limit Theorem: Watch [the entire lesson](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4577183869/e-117708697/m-117640004).
* Lesson 8
  + Confidence intervals: Watch [the entire lesson](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4590858858/m-126004777).
* Lesson 9
  + alpha levels and critical regions: Start at [the beginning of Lesson 9](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4577134255/e-136488372/m-135919217) and continue until you reach the video "Hypotheses".
  + Null and alternative hypotheses and when to reject the null: Start at the video ["Hypotheses"](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4577134255/m-135919254) and continue until you reach the video "Decision Errors".
  + Statistical decision errors: Start at the video ["Decision Errors"](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4577134255/e-136488396/m-135919293) and continue to the end of Lesson 9.
* Lesson 10a
  + One sample t-tests (null and alternative hypotheses, t-statistics, p-values, and when to reject the null): Start at [the beginning of Lesson 10a](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4578095863/e-147019259/m-147019260) and continue until you reach the video "Rent - CI".
  + p-values: Start at the video ["P-Value"](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4578095863/e-147905756/m-147905757) and continue until you reach the video "Rent - t-Critical Values".
  + Confidence intervals for a one-sample t-test: Start at the video ["Rent - CI"](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4578095863/e-147019317/m-147019318) and continue until you reach the video "Dependent Samples".
  + Dependent samples t-tests: Start at the video ["Dependent Samples"](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4578095863/m-147019329) and continue to the end of Lesson 10a.
* Lesson 10b
  + More t-test examples, Cohen's d, and r squared: Watch [the entire lesson](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4621269407/m-230229083).
* Lesson 11   
  + Independent samples t-test: Start at [the beginning of Lesson 11](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4576163829/m-154863474) and continue until you reach the video "Acne Medication".
  + More examples of independent samples t-tests: Start at the video ["Acne Medication"](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4576163829/m-154863508) and continue until you reach the video "Pooled Variance Sum of Squares".
  + Confidence intervals for an independent samples t-test: Start at the video ["Shoes - 95% CI"](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4576163829/e-154863538/m-154863539) and continue until you reach the video "Gender and Shoes".
  + r squared for independent samples t-tests: Watch the video ["Gender and Shoes"](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4576163829/e-154863546/m-154863547).
  + Pooled standard error: Start at the video ["Pooled Variance Sum of Squares"](https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4576163829/e-154863550/m-154863551) and continue to the end of Lesson 11.

**Questions For Investigation**

1. What is our independent variable? What is our dependent variable?
   1. The independent variable is the words condition (congruent / incongruent).
   2. The dependent variable is the time it takes to name the ink colors.
2. What is an appropriate set of hypotheses for this task? What kind of statistical test do you expect to perform? Justify your choices.
   1. Definitions
      1. H0: Null Hypothesis
      2. HA: Alternate Hypothesis
      3. MC: Mean of Congruent Word Condition
      4. MI: Mean of Incongruent Word Condition
   2. H0: Mc - MI = 0
      1. Null Hypothesis: The incongruent word condition did not change the amount of time it took participants to name the color of the words.
      2. Justification: The null hypothesis should always state that there is no change.
   3. HA: Mc - MI < 0
      1. Alternate Hypothesis: It took longer for the participants to name the color of the words with the incongruent word condition versus the congruent word condition.
      2. Justification: This alternate hypothesis is appropriate since I am trying to demonstrate that it takes longer to name the ink colors with the incongruent word condition.
   4. I expect to perform a one-tailed t-test in the negative (-) direction.
      1. Justification: I am trying to demonstrate that it takes longer to name the ink colors with the incongruent word condition, so performing the one-tail test in the negative direction is appropriate. A 95% confidence level and t-critical value of -1.714 will be used to determine if the null hypothesis should be accepted or rejected.

Now it’s your chance to try out the Stroop task for yourself. Go to [this link](https://www.google.com/url?q=https://faculty.washington.edu/chudler/java/ready.html&sa=D&ust=1452645040310000&usg=AFQjCNFZ-cV8yL8k-3IlFq8q2TGHdRnkvQ), which has a Java-based applet for performing the Stroop task. Record the times that you received on the task (you do not need to submit your times to the site.) My times (10.125, 17.772). Now, download [this dataset](https://www.google.com/url?q=https://drive.google.com/file/d/0B9Yf01UaIbUgQXpYb2NhZ29yX1U/view?usp%3Dsharing&sa=D&ust=1452645040310000&usg=AFQjCNGU33sBIMMJ3hhTaPwe_FreR8zPfA) which contains results from a number of participants in the task. Each row of the dataset contains the performance for one participant, with the first number their results on the congruent task and the second number their performance on the incongruent task.

1. Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability.
   1. Congruent word condition (M = 14.051, SD = 3.559).
   2. Incongruent word condition (M = 22.016, SD = 4.797).
   3. Difference of Means (M = -7.965, SD = 4.865).
2. 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.
   1. 
   2. 
   3. Observations:
      1. The Incongruent data appears to be normally distributed. There are also two noticeable outliers to the right.
      2. The congruent data appears to be slightly skewed to the right (positive direction).
      3. In every instance, it took the participants longer to name the ink colors during the incongruent word condition versus the congruent word condition.
3. 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?
   1. Confidence level on the mean difference is 95%, t-critical = -1.714.
   2. Conclusion: We should reject the null hypothesis. It took the participants a significantly longer time (t(23) = -8.021, p < .0001) to complete the exercise under the incongruent work condition (M = 22.016, SD = 4.797) versus the congruent word condition (M = 14.051, SD = 3.559). Yes, the results matched my expectations.
4. 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!

**Rubric:**

**Overview:**

This rubric is here to help you understand the specifications for how your project will be evaluated. It is the same rubric you should share with others who give you feedback. You should look at the rubric **before you begin**working on this project and**before you submit it**.

Before you begin:

1. Read the [final project instructions](https://www.google.com/url?q=https://docs.google.com/document/d/1-OkpZLjG_kX9J6LIQ5IltsqMzVWjh36QpnP2RYpVdPU/pub?embedded%3DTrue&sa=D&ust=1452645577911000&usg=AFQjCNFsjRHGPjiPWzT_ymu28DInRq_spw) and this document in detail.

**Before you submit:**

1. Read the rubric below in detail and do your best to evaluate where your project stands.
2. If you think your project does not meet specifications for **any** criterion, make necessary changes so that it “meets specifications”.
3. When you are confident that your project meets or exceeds specifications in each criterion, share it with others for feedback

**The Rubric:**

|  |  |
| --- | --- |
| Criteria | Meets Specifications |
| Responses to Project Questions |  |
| **Question 1: Identify variables in the experiment** | Question response correctly identifies the independent and dependent variables in the experiment. |
| **Question 2: Establish a hypothesis and statistical test** | Null and alternative hypotheses are clearly stated in words and mathematically. Symbols in the mathematical statement are defined. |
|  | A statistical test has been proposed which will distinguish the proposed hypotheses. Any assumptions made by the statistical test are addressed. |
| **Question 3: Report descriptive statistics** | Descriptive statistics, including at least one measure of centrality and one measure of variability, have been computed for the dataset’s groups. |
| **Question 4: Plot the data** | One or two visualizations have been created that show off the data, including comments on what can be observed in the plot or plots. |
| **Question 5: Perform the statistical test and interpret your results** | A statistical test has been correctly performed and reported, including test statistic, critical test statistic or p-value, and test result. The test results are interpreted in terms of the experimental task performed. |
| **Question 6: Digging deeper and extending the investigation** | Hypotheses regarding the reasons for the effect observed are presented. An extension or related experiment to the performed Stroop task is provided, that may produce similar effects.  ***This question is optional and does not need to be answered in order to meet project specifications.*** |

**References:**

* <https://en.wikipedia.org/wiki/Stroop_effect>
* <https://s3.amazonaws.com/udacity-hosted-downloads/t-table.jpg>
* <http://www.graphpad.com/quickcalcs/pValue2/>