Homework 7: Student's *t*-tests & Making Inferences from Single Samples, Dependent Samples, and Independent Samples

<u>Homework</u>:
Regression Questions 1 - 2
Student's *t*-test Questions 3 - 6

<u>SPSS</u>:

***t* Test for a Single Sample (one-sample *t*-test)**

- Enter the data from Question 3 below into SPSS
- Run a single-sample *t*-test.
- Identify $\overline{X}, S, S_{\overline{X}}, t, df$ and the *p* value on your SPSS output. Circle the numbers and annotate them so that you can identify them later.
- Explain in a sentence or two what these numbers mean in terms of the variables (e.g., IVs and DVs) investigated in the study?
- Compare your results with those you calculated by hand.

***t* Test for Dependent Means**

- Enter the data from Question below 4 into SPSS
- Run a *t*-test for dependent means.
- Identify $\overline{X}_D, S_{\overline{X}_D}, t, df$ and the *p* value on your SPSS output. Circle the numbers and annotate them so that you can identify them later.
- Explain in a sentence or two what these numbers mean in terms of the variables (e.g., IVs and DVs) investigated in the study?
- Compare your results with those you calculated by hand.

***t* Test for Independent Means**

- Enter the data from Question 6 below into SPSS
- Run a *t* test for independent means (with and without Student F's data)
- Identify the significance level of Levene's test for equality of variances, $S_1, S_2, S_{diff}, t, df$ and *p* value on SPSS output. Circle the numbers and annotate them so that you can identify them later. What do these numbers mean in terms of the variables in the study?
- Compare your results with those you calculated by hand. If your means and *t*-test value are not identical, please review your hand calculations and your SPSS data.
- Identify the significance level of the Levene's test for equality of variances. Have you violated any assumptions of the *t*-test?

Regression Review

1. We want to know if the number of cups of tea a person drinks is related to cognitive functioning. Using the model that predicts cognitive functioning from tea drinking, what would cognitive functioning be if someone drank 10 cups of tea? (dataset is "TeaMakesYouBrainy716.sav", which is in your Stats109 data folder).

2. Imagine that a friend is studying the happiness of college students and wants to predict happiness using *multiple* predictors. What advice would you give your friend about choosing predictor variables?

Student's *t*-test

3. A researcher tests five individuals who have seen paid political ads about a particular issue. These people take a multiple-choice test about the issue on which people in general (who know nothing about the issue) usually answer 40 questions correct. The number correct for each of these 5 individuals was 48, 41, 40, 51, and 50. Using alpha of .05 for a two-tailed test, do people who see the ads perform better on this test?
    a. Use the steps of hypothesis testing.
    b. Sketch the distribution and the cutoff scores.

4. A program to decrease littering was carried out in four cities in California's Central Valley starting in August 2007. The amount of litter in the streets (average pounds of litter collected per block per day) was measured during July before the program started and then the next July, after the program had been in effect for a year. The results were as follows:

City	Jul-07	July-08
Fresno	9	2
Merced	10	4
Bakersfield	8	9
Stockton	9	1

    Using the .01 level of significance, was there a significant decrease in the amount of litter?
    a. Use the five steps of hypothesis testing.
    b. Sketch the distribution and the cutoff scores.

5. For each of the following studies, say whether you would use either a *t* test for dependent means or a *t* test for independent means.
    a. A researcher randomly assigns a group of 25 unemployed workers to receive a new job skills program, and then measures how well they all do on a job skills test.
    b. A researcher measures self-esteem in 21 students before and after taking a difficult exam.
    c. A researcher tests reaction time of each member of a group of 14 individuals twice, once while in a very hot room and once while in a normal-temperature room.

6. An educational psychologist was interested in whether using a student's own name in a story affected children's attention span while reading. Six children were assigned randomly to read a story under ordinary conditions (using names like Dick and Jane). Five other children read versions of the same story, but with each child's own name substituted for one of the children in the story. The researcher kept a careful measure of how long it took each child to read the story. The results appear in the following table. *Do not include Student-F's data or otherwise* $N_1 \neq N_2$. *Also, if you look at the graphical distribution (e.g., histogram) of your data, SPSS will adjust the range automatically based on the values in a sample. So just be careful if you ever compare any graphs without looking at the X and Y axes.*

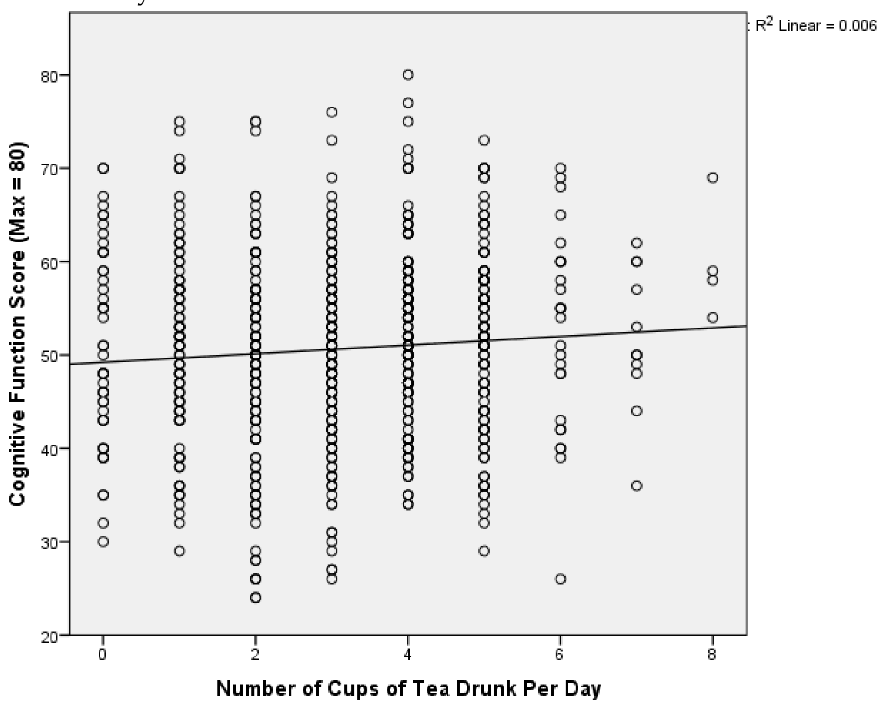
Ordinary story		Own-Name Story	
Student	Reading Time	Student	Reading Time
A	2	G	4
B	5	H	16
C	7	I	11
D	9	J	9
E	6	K	8
F	***7***		

a. Using the .05 level, does including the child's name make any difference? (Use the steps of hypothesis testing.)

b. Sketch the distributions involved.

Homework 7: Student's *t*-tests
Answers

1. **Tea Makes You Brainy**
Step 1: Enter the data into SPSS
Step 2: Create a scatterplot and look at your data. Check that assumptions for regression analysis are met.



Step 3: Run bivariate linear regression in SPSS.

Descriptive Statistics

	Mean	Std. Deviation	N
Cognitive Function Score (Max = 80)	50.61	9.883	716
Number of Cups of Tea Drunk Per Day	3.03	1.669	716

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	49.218	.764		64.382	.000
	Number of Cups of Tea Drunk Per Day	.460	.221	.078	2.081	.038

a. Dependent Variable: Cognitive Function Score (Max = 80)

The *positive standardized beta value* (.078) indicates a positive relationship between number of cups of tea drunk per day and level of cognitive functioning, in that the more tea you drink, the higher your level of cognitive functioning.

Step 4: Use the model to predict level of cognitive functioning after drinking 10 cups of tea.

Cog. Functioning = b_0 (when Tea is 0) + b_1 (Tea Drinking)
= 49.22 + (.460 x Tea Drinking)

Substitute "tea drinking" with 10 cups:
Cog. Functioning = 49.22 + (.460 x 10) = 53.82

Therefore, for drinking 10 cups of tea per day you would predict the cognitive functioning score to be 53.82.

2. **Choosing predictors for multiple regression.**
SPSS will run multiple regression and give you output even if you do not choose appropriate predictors for college student happiness. SPSS does not tell you if the model it creates is valid or generalizable. Choose predictors carefully, based on theoretical rationale and/or past research that has demonstrated that such variables would be valid predictors for your research question. The relationship between predictors and your criterion should be linear. Also be sure to pay attention to how strongly correlated the predictor variables are to each other (collinearity and multicollinearity). High multicollinearity can lead to untrustworthy beta coefficients, the variables may not account for much unique variance, and you may have difficulty determining the importance of predictor variables.

(Bonus: You can check for multicollinearity using the VIF/tolerance statistic. VIF greater than 10 or below 0.2 shows that there may be potential problems. You can also use descriptive statistics to check the correlation matrix for multicollinearity. For more details see homework 6, class notes, and Field, Section 8.5, p. 321-326).

3. Single Sample *t*-test

$N = 5; \mu = 40$

$df = N - 1 = 5 - 1 = 4$

$\alpha = .05_{2-tailed}$

Individual	X	$X - \bar{X}$	$\left(X - \bar{X}\right)^2$
1	48	2	4
2	41	-5	25
3	40	-6	36
4	51	5	25
5	50	4	16
Total =	230		SS=106
Sample Mean =	$\frac{230}{5} = 46$		

1. Population 1: People who have seen paid political ads.
   Population 2: People in general.

Null hypothesis: $\mu_1 = \mu_2$; Alternative (Research) hypothesis: $\mu_1 \neq \mu_2$
Note: Given the two-tailed nature of this test, this question should probably ask – "Do people who see political ads perform *differently* from those who do not see such ads?"

2. Comparison distribution: *t*-distribution; pop. variance is not known; $\mu_{\bar{X}} = \mu = 40$

Estimated population variance: $S^2 = \frac{SS}{N-1} = \frac{SS}{df} = \frac{106}{4} = 26.5$

Estimated population standard deviation: $\sqrt{S^2} = \sqrt{26.5} = 5.148$

Estimated Standard deviation of the distribution of means (standard error):

$$S_{\bar{X}} = \frac{S}{\sqrt{N}} = \frac{5.148}{\sqrt{5}} = 2.24$$

Shape: *t* distribution when $df = 4$ (not normal, more leptokurtic because it's a *t* dist.)

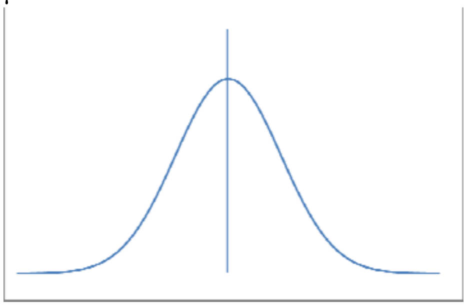
3. Cutoff (critical) scores for a 2-tailed *t*-test at the $\sigma = .05$ level for *df* = 4 are -2.776 and +2.776 (See *t*-table).

4. $t_{\alpha=.05} = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} = \frac{46 - 40}{2.24} = 2.68 \quad \therefore \quad t(4) = 2.68$

5. A critical or cutoff *t* value of 2.68 is not more extreme than ±2.776, so we fail to reject the null hypothesis that people who have seen paid political ads about a particular issue score the same as people in general. The results of the study indicate that ads likely do not matter. **Note**: Keep in mind that only 5 people were included in the study.

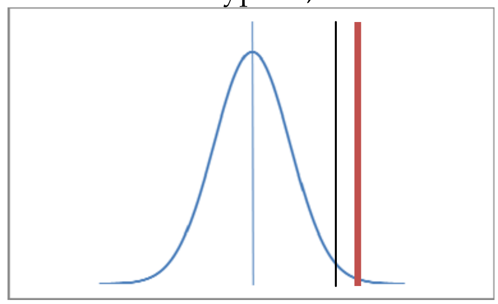
Sketch of the distribution and the cutoff scores.

Population Distribution (Normal);
μ=40



The normal distribution of the pop.

Comparison Distribution (*t*)
Under the null hypoth., the mean = 0



We compare our calculated *t-value* (2.60) to the critical *t*-value from the *t*-table (±2.776).

4. Dependent/Paired Samples *t*-test

$N = 4$

$\bar{X}_{difference} = -5($

$df = N - 1 = 4 - 1 = 3$

$\alpha_{(\text{one}-tailed)} = .01$

City	Jul-07	July-08	Difference	Diff – Mean Diff	(Diff – Mean Diff) ^{2}
Fresno	9	2	-7	(-7)–(-5) = -2	$2^2 = 4$
Merced	10	4	-6	-1	1
Bakersfield	8	9	1	6	36
Stockton	9	1	-8	-3	9
		Total =	-20		SS_{Diff} = 50
		Mean (-20/4) =	-5		

a.

1. Population 1: Cities that show no change from July-07 to July-08 (no change = 0)
   Population 2: Cities that ran the program.

Therefore, Null hypothesis: $\mu_1 = \mu_2$; Alternative (Research) hypothesis: $\mu_1 \neq \mu_2$

2. Comparison distribution: *t*-distribution because pop. parameters (e.g., μ, σ) are *not known*; there are also two samples. Under the null hypothesis, both population means should be the same. So on average difference between sample means would theoretically be 0; $\mu_{\bar{X}_D} = 0$

Estimated population *variance* (estimated from sample data): $S_D^2 = \frac{SS_D}{df} = \frac{50}{3} = 16.67$

Estimated standard *deviation* (estimated from sample data): $S_D = \sqrt{S_D^2} = \sqrt{16.67} = 4.08$

Estimated standard deviation of the sampling distribution of mean differences (standard error of the mean of *difference scores*): $S_{\bar{X}_D}$ or $S_{\bar{D}} = \frac{S_D}{\sqrt{N}} = \frac{4.08}{\sqrt{4}} = 2.04$

Shape = *t* distribution when $df = 3$.

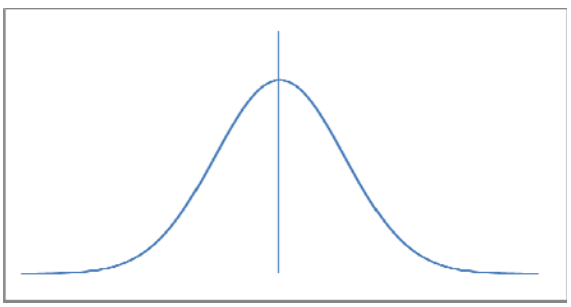
3. Cutoff for a two-tailed *t*-test with $\alpha = .01$ level for $df = 3$ is ±5.84.

4. $t_{\bar{X}_D \alpha=.01} = \frac{\bar{X}_D - \mu_0}{S_{\bar{X}_D}} = \frac{\bar{X}_D - 0}{S_{\bar{X}_D}} = \frac{\bar{X}_D}{S_{\bar{X}_D}} = \frac{-5}{2.04} = -2.45 \quad \therefore t(df) = ?;\ t(3) = -2.45$

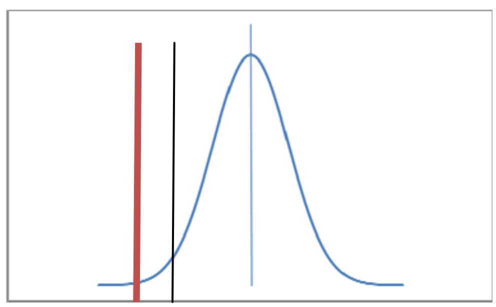
7. The calculated/obtained *t-value* of -2.45 is **not** more extreme than the critical *t-value* of ±5.84, so we fail to reject the null hypothesis. We retain that there was no significant decrease in littering for those cities between July-07 to July-08. In other words, the difference is not greater than we might expect due to chance/random error.

Sketch of the distribution and the cutoff scores.

Population of Difference Scores
μ=0



Comparison Distribution (t)
Under the null hypoth., the mean = 0



We compare our calculated *t-value* (-2.45) to the critical-*t* value from the *t*-table (±5.84).

5. **Dependent or Independent t-Tests?**
   - **a.** Randomly assigns 25 workers to receive a job skills program, measures job skill test outcome. Use *t*-test for **independent** means.
   - **b.** Researcher measures self-esteem in 21 people before and after an exam. Use *t*-test for **dependent** means.
   - **c.** Researcher measures reaction time for 14 individuals under two different conditions. Use *t*-test for **dependent** means.

6. Independent Sample *t*-test

t-test for independent means α = .05, two-tailed test

Ordinary story	
X	$\left(X-\bar{X}\right)^2$
2	14.44
5	0.64
7	1.44
9	10.24
6	0.04

$\bar{X}_1 =$ 5.8 $SS_1 =$ 26.80
N_1 =5
df_1=4

Own-Name Story	
X	$\left(X-\bar{X}\right)^2$
4	31.36
16	40.96
11	1.96
9	0.36
8	2.56

$\bar{X}_2 =$ 9.6 $SS_2 =$ 77.20
N_2 =5
df_2=4

a.
1. Population 1: Children who read a story with their own name.
   Population 2: Children who read a story under ordinary conditions.

   Null hypothesis: $\mu_1 = \mu_2$
   Alternative (Research) hypothesis: $\mu_1 \neq \mu_2$

2. Comparison distribution:

- **Mean of comparison distribution:**
  Distribution of *differences between means* $\mu_{\bar{X}_1-\bar{X}_2} = 0$

- **Estimated population variances:**

$$S_1^2 = \frac{SS_1}{N_1-1} = \frac{26.8}{4} = 6.7$$

$$S_2^2 = \frac{SS_2}{N_2-1} = \frac{77.2}{4} = 19.3$$

- **Pooled estimate of population variance:**

$$S^2_{pooled} \ or \ S^2_p = \frac{SS_1 + SS_2}{df_1 + df_2}$$

$$= \frac{26.8 + 77.2}{4 + 4} = \frac{104}{8} = 13$$

OR if using the weighted approach that is appropriate for unequal sample sizes:

$df_{total} = df_1 + df_2 = 4 + 4 = 8$

$$S^2_{Pooled} = \frac{df_1}{df_{Total}}(S_1{}^2) + \frac{df_2}{df_{Total}}(S_2{}^2)$$

$$= \frac{4}{8}(6.7) + \frac{4}{8}(19.3)$$

$$= 3.35 + 9.65$$

$$= 13$$

- **Variance of each mean distribution:**

$$S^2_{\bar{X}_1} = \frac{S^2_P}{N_1} = \frac{13}{5} = 2.6$$

$$S^2_{\bar{X}_2} = \frac{S^2_P}{N_2} = \frac{13}{5} = 2.6$$

- **Variance of distribution of differences between means:**

$$S^2_{Difference} = S^2_{\bar{X}_1} + S^2_{\bar{X}_2} = 2.6 + 2.6 = 5.2$$

- **Standard error of differences/Standard deviation of distribution of differences between means:**

$$S_{diff.} = \sqrt{S^2_{Difference}} = \sqrt{5.2} = 2.28$$

But, because samples are of equal size, we can skip the calculation of variance of differences between means and just calculate S_{diff} based on the pooled variance.

$$S_{diff.} = \sqrt{\frac{S^2_p}{N_1} + \frac{S^2_p}{N_2}}$$

$$= \sqrt{\frac{13}{5} + \frac{13}{5}}$$

$$= \sqrt{2.6 + 2.6}$$

$$= \sqrt{5.2}$$

$$= 2.28$$

- **Shape of comparison distribution**: *t* distribution with *df* = total = 8

3. When $df = 8$, the cutoff/critical $t_{two-tailed-\alpha=.05} = \pm 2.306$

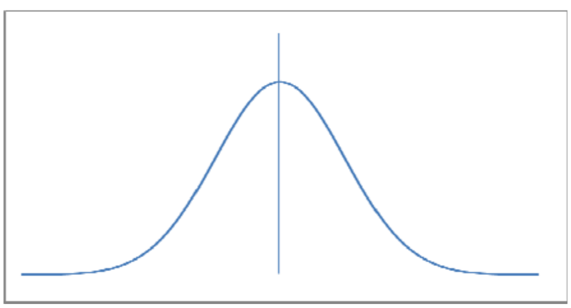
4. $t = \dfrac{\bar{X}_1 - \bar{X}_2}{S_{Difference}} = \dfrac{5.8 - 9.6}{2.28} = -1.6667 \quad \therefore t(df) = ?; \; t(8) = -1.667$

NOTE: Always look at your means! What does a negative *t*-value represent? The *t*-value can be either positive or negative depending on which sample you determine is #1 and which is #2.

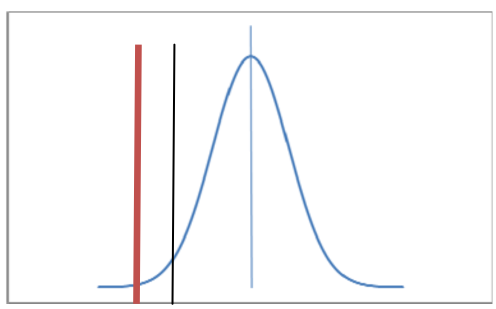
5. The calculated/obtained *t-value* of -1.667 is **not** more extreme than the critical *t-value* of -2.306, so we fail to reject the null hypothesis. The difference between the means is not large enough (based on the standard error of 2.28) to determine that the difference unlikely to occur due to chance alone. The difference in means might just reflect random variability in scores when $df = 8$. Thus, no statistical difference exists between reading scores for those children who read a story containing their own names and those who read a story under ordinary conditions.

Sketch of the distribution and the cutoff scores.

Population of Mean Differences
μ=0



Comparison Distribution (t)
Under the null hypoth., the mean = 0



We compare our calculated *t-value* (-1.667) to the critical-*t* value from the *t*-table (±2.306).

t-Test for a Single Sample/One-Sample t-test

Syntax:

```
T-TEST
 /TESTVAL = 40
 /MISSING=ANALYSIS
 /VARIABLES=CorrectAnswers
 /CRITERIA=CI(.95).
```

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Correct Answers	5	46.00	5.148	2.302

Sample mean

Estimated Std. Dev. (S)

Std. Error of the mean $\left(S_{\bar{X}}\right)$

One-Sample Test

	Test Value = 40					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference Lower	Upper
Correct Answers	2.606	4	.060	6.000	-.39	12.39

Obtained t = 2.606
Sign of t tells you direction of difference.

Degrees of Freedom = 4

Sig. 2 tailed: p = 0.06; compare p to alpha

IV = Political tv ad (yes or no), and DV = number of correct answers on a multiple choice test. We want to test if people who see the ads do better on this multiple choice test (using the .05 level of significance, two-tailed).

Average number of correct answers with ad = 46 (see X bar above)
The one-sample t-test compares the sample value of 46 (sample mean) to the "test value" of 40 (the population mean) and determines how far from the population mean the sample mean is in terms of standard error units (S). The statistic is $t(4)$ = 2.606 indicates the test group has a greater value in the DV (mean difference of that value is 6.0).
df = 4 (5 people - 1 person's score)
p =0.060, not significant if alpha is .05. We fail to reject the null hypothesis that there is a difference between the scores of people who see ads and people who do not see ads.

t-Test for a Dependent Means

Syntax:

```
T-TEST PAIRS=July07 WITH July08 (PAIRED)
  /CRITERIA=CI(.9500)
  /MISSING=ANALYSIS.
```

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	July07	9.00	4	.816	.408
	July08	4.00	4	3.559	1.780

$\bar{X}$ for each year

$S_{\bar{X}_D} = 4.082$

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	95% Confidence Interval of the Difference Upper			
Pair 1 July07 - July08	5.000	4.082	2.041	-1.496	11.496	2.449	3	.092

$\bar{X}_D = 5$

$t = 2.449$

$df = 3$

$p = 0.092$
NOT significant

Identify $\bar{X}_D, S_{\bar{X}_D}, t, df$ and the p value on your output. Circle the numbers and annotate.

IV = year, with two levels
DV = amount of litter in the streets (average pounds of litter collected per block per day)
Was there a change in litter over time?
There was no statistically significant decrease in litter in the streets over the course of one year.
t = 2.449 (There was a positive mean difference between the two dates of data collection)
df = 3
p = 0.092; fail to reject null hypothesis that there is no significant difference in litter on the streets between cities that use the program and cities that do not use the program.

SPSS

***t*-Test for Independent Means Without Subject/Participant F's data**

Syntax:

```
T-TEST GROUPS=Story Type (1 2)
 /MISSING=ANALYSIS
 /VARIABLES=Reading Time
 /CRITERIA=CI(.95).
```

Group Statistics

	Story Type	N	Mean	Std. Deviation	Std. Error Mean
Reading Time	Ordinary Story	5	5.80	2.588	1.158
	Own Name Story	5	9.60	4.393	1.965

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
ReadingTime	Equal variances assumed	.853	.383	-1.666	8	.134	-3.800	2.280	-9.058	1.458
	Equal variances not assumed			-1.666	6.479	.143	-3.800	2.280	-9.281	1.681

Levene's Test
p = 0.383

t= -1.66

df= 8

p = 0.134

S_{diff} = 2.280; standard deviation of comparison distribution

Identify the significance level of Levene's test for equality of variances, $S_1, S_2, S_{diff}, t, df$ and p

IV: Ordinary condition or own name condition
DV: length of time spent reading
t = -1.66 (Own name story condition has longer reading times)
df = 8 (4 *df* from each condition)
p = 0.134; fail to reject null hypothesis that there is no significant difference between reading conditions.

t Test for Independent Means With Subject F's data

Syntax:

```
T-TEST GROUPS=Story Type (1 2)
  /MISSING=ANALYSIS
  /VARIABLES=Reading Time
  /CRITERIA=CI (.95).
```

Group Statistics

	Story Type	N	Mean	Std. Deviation	Std. Error Mean
Reading Time	Ordinary Story	6	6.00	2.366	.966
	Own Name Story	5	9.60	4.393	1.965

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
ReadingTime	Equal variances assumed	1.301	.284	-1.739	9	.116	-3.600	2.070	-8.283	1.083
	Equal variances not assumed			-1.644	5.893	.152	-3.600	2.189	-8.981	1.781

df = 9

S_{diff} = 2.070
standard deviation of comparison distribution

Levene's Test
p = 0.383

t = -1.739

p = 0.116

Identify the significance level of Levene's test for equality of variances, $S_1, S_2, S_{diff}, t, df$ and p

IV: Ordinary condition or own name condition
DV: length of time spent reading
t = -1.66 (Own name story condition has *numerically longer* reading times; but see p value)
df = 9 (5 df from ordinary story condition, 4 df from own name story condition)
p = 0.116; fail to reject null hypothesis that there is no significant difference between reading conditions. Conclude that there are no statistically significant differences in reading time