A08 Statistics Problems

Name: Max Hodgen **Grade = /24**

1. A high school asked its employees to estimate the how much time per week they required their students to use the Internet for instructional purposes. The following chart shows the times estimated by each teacher and then divided into two groups by the gender of the teacher.

Estimated Time in minutes per week

Female	Male
15	5
75	15
5	10
60	80
20	10
40	0
30	10
15	10
45	15
40	20
20	10
35	20
15	15
80	10
10	0
10	0
45	15
25	5
10	10
65	15

a. Run the appropriate t-test on the data and explain why you chose this form of t-test. (3 points)

I chose to use a 2 sample t-test because we wanted to compare the internet usage of high school students between two different groups (male & female teachers).

b. paste a copy of the Session Window in the box below: (1 point)

```
Two-sample T for Female vs Male

         N  Mean  StDev  SE Mean
Female  20  33.0   22.8      5.1
Male    20  13.8   16.7      3.7


Difference = mu (Female) - mu (Male)
Estimate for difference:  19.25
95% CI for difference:  (6.41, 32.09)
T-Test of difference = 0 (vs not =): T-Value = 3.05  P-Value = 0.004  DF = 34
```

c. Explain what the results mean for these two groups. (3 points)

The mean value for the time (minutes) required for internet usage of students for the female teachers was 33 minutes compared to 13.8 minutes for the male teachers. To see if there was a statistical significant difference between the groups a two sample t-test was preformed resulting in a P-value of 0.004. The low p-value indicates that there is only a 0.4 % chance that the differences in means was caused by natural variation or that there is a statistical significant difference between the required amount of student internet usage between male and female teachers.

Can we say this? Since the p-value is less than 5% then we can reject the null hypothesis, which is the antithesis of the researcher's hypothesis.

2. Jim Thompson's 8th grade science class took the pretest on the solar system and a then posttest after a week of class work learning the content.

Test Scores on Solar System

Student	Pretest	Posttest
A	36	42
B	53	55
C	41	52
D	34	50
E	50	48
F	53	49
G	48	50
H	20	21
I	42	52
J	39	52
K	45	50
L	55	56
M	46	48
N	50	48
O	41	53
P	45	28
Q	42	52
R	25	32
S	45	42
T	43	50
U	20	25
V	74	77
W	50	48
X	46	50

a. Run the appropriate t-test on the data and explain why you chose this form of t-test. (3 points)

I chose to run a paired sample t-test to compare the pretest and posttest of the same individual in the group.

b. Paste a copy of the Session Window in the box below: (1 point)

```
Paired T for Pretest - Posttest

             N   Mean  StDev  SE Mean
Pretest     24  43.46  11.50     2.35
Posttest    24  47.08  11.48     2.34
Difference  24  -3.62   6.99     1.43


95% CI for mean difference: (-6.58, -0.67)
T-Test of mean difference = 0 (vs not = 0): T-Value = -2.54  P-Value = 0.018
```

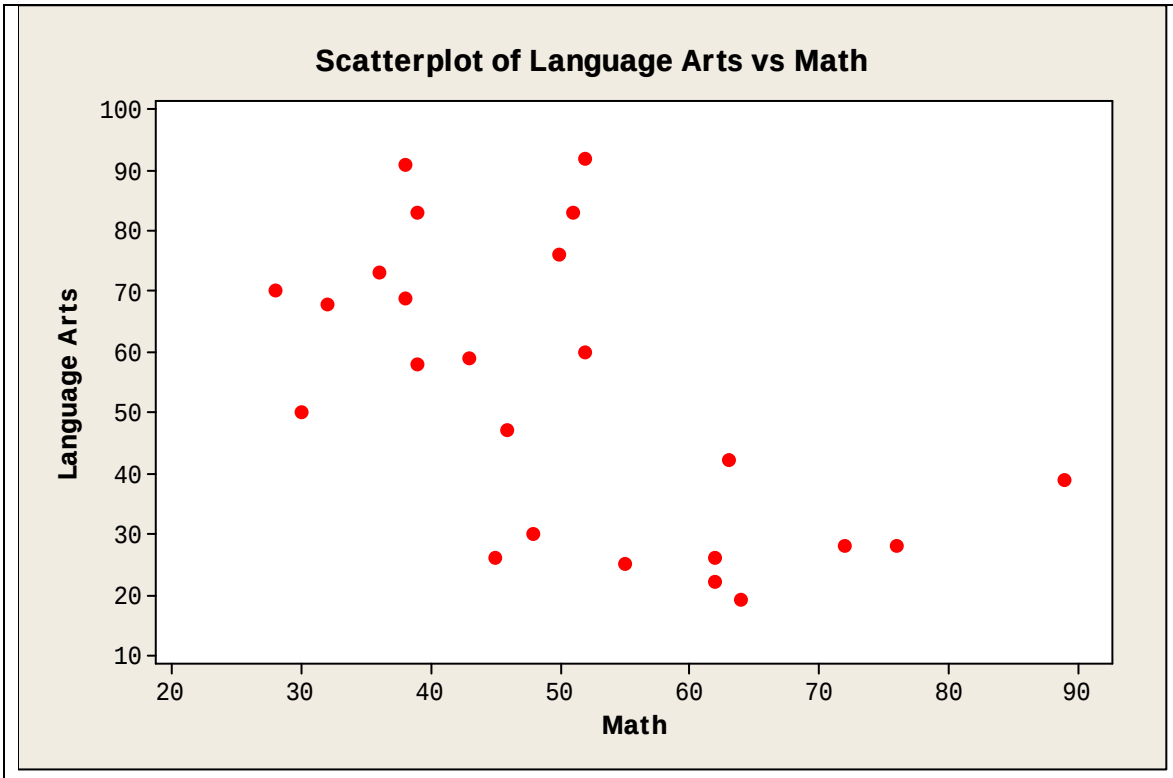
c. Explain what the results mean for the class. (3 points)

Analyzing the means shows an increase of 3.62 for the posttest when compared to the pretest. To see if there a statistical significant difference a paired sample t-test was performed resulting in a P-value =0.018. That means there is only a 1.8% chance that the difference in means was caused by natural variation or since the p-value was less than 5% (alpha level or risk level) that there is a statistical significant difference between the pretest and the posttest.

3. Senior students in a small high school class took a national standardized test required for admission to college. The results, in percentile ranks, are shown below: (10 points total)

Student	Language Arts	Math
A	28	76
B	50	30
C	39	89
D	59	43
E	30	48
F	73	36
G	91	38
H	42	63
I	76	50
J	26	45
K	58	39
L	19	64
M	83	39
N	70	28
O	69	38
P	22	62
Q	26	62
R	83	51
S	25	55
T	47	46
U	28	72
V	68	32
W	60	52
X	92	52

a. Using ***all*** appropriate statistics covered in the class, to date, use Minitab to analyze each column of data as well as to analyze relationships and differences between the two standardized tests. Paste copies of the *different* Session Windows you used in your analyses in the box below:



Descriptive Statistics: Math, Language Arts

Variable	N	N*	Mean	SE Mean	TrMean	StDev	Minimum	Q1	Median
Math	24	0	50.42	3.13	49.68	15.31	28.00	38.25	49.00
Language Arts	24	0	52.67	4.87	52.41	23.84	19.00	28.00	54.00

Variable	Q3	Maximum
Math	62.00	89.00
Language Arts	72.25	92.00

Correlations: Math, Language Arts

Pearson correlation of Math and Language Arts = -0.573
P-Value = 0.003

Two-sample T for Language Arts vs Math

	N	Mean	StDev	SE Mean
Language Arts	24	52.7	23.8	4.9
Math	24	50.4	15.3	3.1

Difference = mu (Language Arts) - mu (Math)
Estimate for difference: 2.25
95% CI for difference: (-9.45, 13.95)
T-Test of difference = 0 (vs not =): T-Value = 0.39 P-Value = 0.699 DF = 39

Paired T for Language Arts - Math

	N	Mean	StDev	SE Mean
Language Arts	24	52.67	23.84	4.87
Math	24	50.42	15.31	3.13
Difference	24	2.25	34.94	7.13

95% CI for mean difference: (-12.50, 17.00)

```
T-Test of mean difference = 0 (vs not = 0): T-Value = 0.32  P-Value = 0.755
```

b. For each of the statistics analyses you did, explain the meaning of the results.

The descriptive statistics for the math group show a mean (50.42) which is slightly higher than the median (49). Therefore, the average of the top half of the data set is farther away from the median than the average of lower half of data. The opposite is true for the language arts group. The trimmed means for both tests were very close to the means but still slightly lower in both cases, indicating that the upper extreme for both tests was farther away from the mean than the lower extreme. Since the standard deviation of the math tests were greater than the language art's tests we can conclude that the distribution of scores was more spread out for the math scores.

The scatter plot of the data indicates a slight inverse trend between math scores and language arts scores. The correlation coefficient confirms the initial trends of the scatter plot because an r-value = -0.573 means there is moderate inverse relationship between the two groups.

A paired t-test was used to compare the performance of the same group of students on math and language arts tests. A p-value of 0.755 indicates that there is a 75.5% chance that the difference in the means was caused by expected variation, or that there was no statistical significant difference between the means of the math and language arts scores.