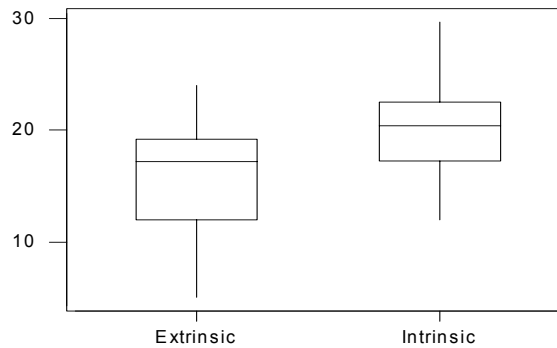


I. Motivation and Creativity – A Randomized Experiment

(From The Statistical Sleuth by Ramsey and Schafer)

Intrinsic	Extrinsic
12.0	5.0
12.0	5.4
12.9	6.1
13.6	10.9
16.6	11.8
17.2	12.0
17.5	12.3
18.2	14.8
19.1	15.0
19.3	16.8
19.8	17.2
20.3	17.2
20.5	17.4
20.6	17.5
21.3	18.5
21.6	18.7
22.1	18.7
22.2	19.2
22.6	19.5
23.1	20.7
24.0	21.2
24.3	22.1
26.7	24.0
29.7	



Descriptive Statistics: Intrinsic, Extrinsic

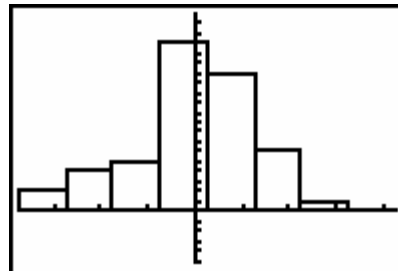
Variable	N	Mean	Median	TrMean	StDev	SE Mean
Intrinsic	24	19.883	20.400	19.795	4.440	0.906
Extrinsic	23	15.74	17.20	15.86	5.25	1.10

Variable	Minimum	Maximum	Q1	Q3
Intrinsic	12.000	29.700	17.275	22.500
Extrinsic	5.00	24.00	12.00	19.20

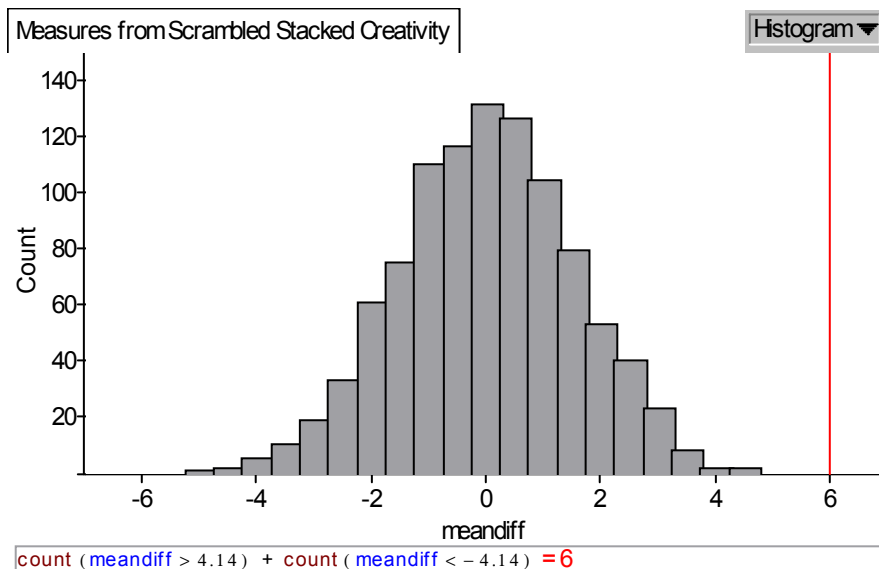
Consider the following question: is there evidence of a treatment effect? If not, then each subject would have received precisely the same creativity score whether (s)he was in the intrinsic or extrinsic group. In that case, we can examine the difference in the mean scores for the two treatment groups if the random assignment process had turned out differently. The TI-83 program RNDMTST allows us to simulate the approximate *randomization distribution* of $\bar{x}_{ext} - \bar{x}_{intr}$. The histogram below displays the values of $\bar{x}_{ext} - \bar{x}_{intr}$ for 50 trials of a simulation. In 50 trials, none of the values of $\bar{x}_{ext} - \bar{x}_{intr}$ exceeded the observed difference, -4.14.

```

PRGM RNDMTST
NO TRIALS
N=?
  
```



Here are the results of a Fathom simulation of the approximate randomization distribution.



Finally, here are the results from a two-sample *t* test.

```

2-SampTTest
μ1≠μ2
t=-2.915292
P=.0056175338
df=43.10803611
x̄1=15.73913043
↓x̄2=19.88333333
  
```

So, after all is said and done, what have we learned here?

II. Predicting Eruptions of Old Faithful



While on a tour of the National Parks with my wife and two stepsons, I had the whimsical idea one morning to reroute our itinerary through Yellowstone. I had never been there, and was particularly interested in watching an eruption of the Old Faithful geyser. I'll spare you the gory details of the six-hour, round-trip drive that was necessary to accomplish my objective. When we arrived, our first thought was "How long will it be until Old Faithful erupts?" We scurried over to the Visitor Center to get our bearings, and to inquire about the projected wait. The Park Rangers had obviously anticipated the question, because there was a display inside that gave fairly precise information about the timing of Old Faithful's eruptions. Not completely satisfied with how these predictions came to be, I set out to collect some data to answer the question for myself. Here's some of what I have discovered.

- Prior to the Borah Peak earthquake of 1983, Yellowstone National Park predicted eruptions of Old Faithful by the equation $I = 30 + 10D$, where
 - I = interval between the current eruption and the next eruption
 - D = duration (in minutes) of continuous visible water on the current eruption.

So the amount of time between the current eruption and the next eruption (I) was predicted to be 10 times the duration of the current eruption (in minutes) plus 30 minutes.

- Since the 1983 earthquake, the Yellowstone National Park has predicted the interval between the current eruption and the next eruption of Old Faithful via another equation. The predictions given in the [brochure](#) "You Too Can Predict Old Faithful" are based on that equation.

The following is a reproduction of the "You Too Can Predict Old Faithful" brochure written by the Old Faithful Staff, Yellowstone National Park, revised 5/95:

YOU TOO CAN PREDICT OLD FAITHFUL

In 1938, a ranger discovered a correlation between Old Faithful Geyser's duration and its subsequent interval. While the specific numbers have changed over time, we still use this principle to predict Old Faithful's eruptions today. You can too! Here's how:

1. Take note of the time that the water comes out and stays out of the cone (the beginning of the continuous water phase).
2. Time the length of the eruption until absolutely no more water is splashing out of the cone.
3. If the eruption lasts: To the starting time add:
 - 1.5 minutes 51 minutes
 - 2.0 minutes 58 minutes
 - 2.5 minutes 65 minutes
 - 3.0 minutes 71 minutes
 - 3.5 minutes 76 minutes
 - 4.0 minutes 82 minutes
 - 4.5 minutes 89 minutes
 - 5.0 minutes 95 minutes

As you can see, the time between eruptions is based on the length of the eruption.

- The brochure "Facts and Theories about Old Faithful Geyser" (revised 3/91) from Yellowstone National Park provides the following basic information about Old Faithful eruption heights and time intervals between eruptions:

	Height (feet)	Interval (minutes)
Minimum	106	41
Average	130	76.17
Maximum	184	114

The values in the table above were obtained from 6,900 eruptions in 1988.

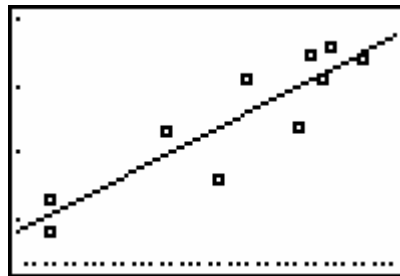
Now we can begin to ask (and answer!) some questions about the eruptions of Old Faithful.

Data from 1978/1979

Run the program OLDFAITH on your TI-83. Enter the values shown for M and N, then watch what happens. Do you know what it's doing?

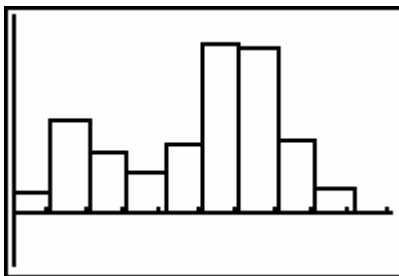
```

PrgmOLDFAITH
HOW MANY SAMPLE...
M=?5
SAMPLE SIZE?
N=?10
  
```

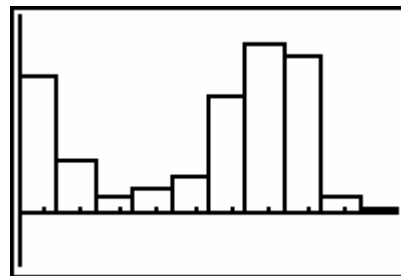


The data that has now been entered into lists L1 and L2 hold the intervals between eruptions of "Old Faithful" and the duration of each eruption during periods of 1978 and 1979. Let's begin by examining the variables individually – the **data analysis** portion of the problem.

Interval



Duration



```

1-Var Stats
x=71.009000901
Σx=15764
Σx²=1155590
Sx=12.79917665
σx=12.77031714
↓n=222
  
```

```

1-Var Stats
x=3.576126126
Σx=793.9
Σx²=3098.75
Sx=1.083950101
σx=1.081506016
↓n=222
  
```

```

1-Var Stats
↑n=222
minX=42
Q1=60
Med=75
Q3=81
maxX=95

```

```

1-Var Stats
↑n=222
minX=1.7
Q1=2.3
Med=4
Q3=4.4
maxX=5.2

```

- Q1:** What do the graphs and numerical summaries tell us about eruptions of Old Faithful?
- Q2:** Has the interval between eruptions changed between 1978/79 and 1988?

This is our first inference question concerning eruptions of Old Faithful. Let's take a close look at precisely what this question is asking before we decide how to proceed.

Here are the results of two statistical inference procedures from the TI-83.

```

TInterval
(69.316,72.702)
x̄=71.00900901
Sx=12.79917665
n=222

```

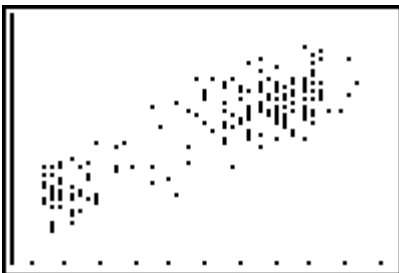
```

T-Test
μ≠76.17
t=-6.007967227
P=7.6602401E-9
x̄=71.00900901
Sx=12.79917665
n=222

```

- Q3:** Is there a relationship between the length of the previous eruption of Old Faithful and the length of time until the next eruption?

Here is a scatterplot of the data.



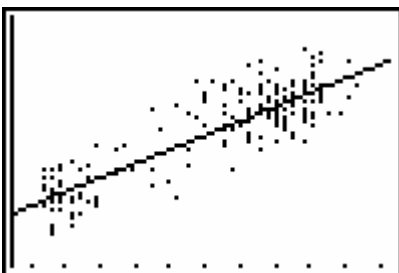
Calculating a least squares regression model:

```

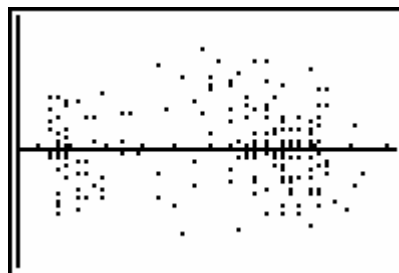
LinReg
y=a+bx
a=33.96675721
b=10.3582062
r²=.7695265669
r=.8772266337

```

How well does the model fit the data?



Examining a residual plot for this model:

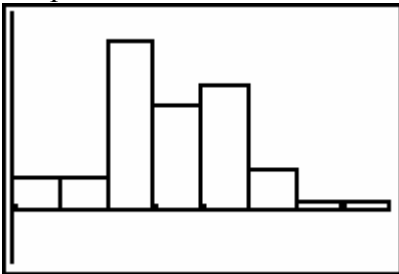


Q4: What does the OLDFAITH program do, anyway?

IDEA: If we treat these data as the “population” of data that describe this relationship, and the linear regression model as the “true” population regression line, we can study the sampling distribution of b , the slope of the sample regression line.

METHOD: Take 50 samples of size $n = 10$ from this population. For each sample, compute the sample regression line and record the slope, b .

Graph of the 50 b -values.

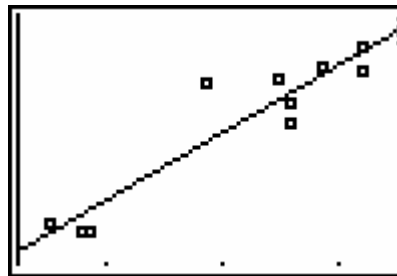


Numerical summaries.

```
1-Var Stats
x̄=10.34980246
Σx=517.4901228
Σx²=5480.96493
Sx=1.59747497
σx=1.581419538
↓n=50
```

PROGRAMS: OLDFAITH and OFQUIK

```
Pr9mOLDFAITH
HOW MANY SAMPLE...
M=?50
SAMPLE SIZE?
N=?10
```



```
Pr9mOLDFAITH
HOW MANY SAMPLE...
M=?1
SAMPLE SIZE?
N=?10
SLOPE
12.76657553
```

IDEA 2: But what if we view the original data as a sample from a larger population of eruptions of the Old Faithful geyser?

Q5: Is there evidence of a statistically significant relationship between interval and duration of Old Faithful eruptions?

```
LinRegTTest
y=a+bx
β≠0 and ρ≠0
t=27.10270973
p=4.729718E-72
df=220
↓a=33.96675721
```

```
LinRegTTest
y=a+bx
β≠0 and ρ≠0
↑b=10.3582062
s=6.15853289
r²=.7695265669
r=.8772266337
```

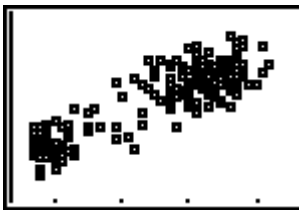
Can we perform inference about the slope, β , of the population regression line, $y = \alpha + \beta x$? Let's examine the required conditions.

- ❑ **The data are an SRS from the population of interest (observations are independent).**

We do not know exactly how these data were obtained, so we may not be able to view them as an SRS from the population of all eruptions of Old Faithful during a set time period. If the sample does not represent the population well, then any inference we perform will be in vain.

- ❑ **The true relationship between duration and interval has a linear form.**

The scatterplot of the sample data certainly appears to have a fairly strong, positive, linear association. (We saw earlier that $r = 0.877$.) A plot of residuals from the least squares regression line, $\hat{y} = 33.967 + 10.358x$, is shown at the right below. The residuals appear to be fairly randomly scattered, indicating that the linear model is an appropriate description of the relationship.

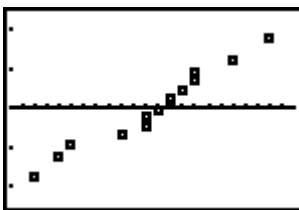


- ❑ **The standard deviation of the response about the population regression line is the same for all values of the explanatory variable.**

The residual plot shows pretty equal scatter at most x -values, though there is a little more scatter for x -values near the center of the plot and a little less scatter for x -values near the edges of the plot.

- ❑ **The response varies normally about the population regression line.**

There are enough data points to examine this condition for some specific x -values. For example, here is a normal probability plot of the values of the response variable for each data point with $x = 1.8$. The strong linear form shown in this plot suggests that the response varies normally in the subpopulation for which $x = 1.8$.



Now that we have examined the conditions for inference about the regression slope, we can construct a 95% confidence interval for β :

$$b \pm t^* \cdot \frac{s}{\sqrt{\sum (resid)^2}} = 10.358 \pm 1.984 \cdot \frac{6.159}{\sqrt{8344}} = 10.358 \pm 0.134 = (10.224, 10.492).$$

CONCLUSION: We can now see that for every minute longer an eruption of Old Faithful lasts, we can expect that it will take an additional 10.224 to 10.492 minutes for the next eruption to occur.

III. Did Mr. Starnes Stack His Class?

Mr. Starnes, in addition to his duties as a teacher of AP Statistics, also creates students' schedules. The other AP Statistics instructor suspects that Mr. Starnes may not have assigned students at random to the two classes. Said less politely, she wonders if Mr. Starnes has loaded his class with brighter students. Instead of challenging her to a duel, Mr. Starnes decides to allow you, an independent statistical consultant, to rule on this matter.

The two principals agreed that SAT Math score and cumulative grade point average are the best indicators of students' abilities. Here are the data.

Starnes GPA	McGrail GPA	Starnes SAT-M	McGrail SAT-M
2.9	2.9	670	620
2.86	3.3	520	590
2.6	3.98	570	650
3.6	2.9	710	600
3.2	3.2	600	620
2.7	3.5	590	680
3.1	2.8	640	500
3.085	2.9	570	502.5
3.75	3.95	710	640
3.4	3.1	630	630
3.338	2.85	630	580
3.56	2.9	670	590
3.8	3.245	650	600
3.2	3.0	660	600
3.1	3.0	510	620
	2.8		580
	2.9		600
	3.2		600

Did Mr. Starnes stack his class? Give appropriate statistical evidence to support your conclusion.