Exploring *t* - Tests

In the reality of experimentation and performing inference tests, there are rare conditions where we know the population standard deviation. Most of the time, we need to use the standard error from properly gathered data. Not knowing the population standard deviation, or σ, means we cannot use the *z*-test with proper accuracy. We need another type of statistical test and this where the *t*-test and *t* confidence interval come in to help us. These procedures allow us to use the sample standard deviation along with **degrees of freedom** to draw conclusions concerning the population of interest.
Degrees of freedom are centered about using $n - 1$ where n is the sample size.

[11]Question: Are women getting the recommended daily allowance of calcium, 1200mg?

Estimates of the daily intake of calcium (in milligrams) for 38 women between the ages of 18 and 24 were taken and are displayed in the table below.

Daily Intake of Calcium (mg)					
808	882	1062	970	909	802
374	416	784	997	651	716
438	1420	1425	948	1050	976
572	403	626	774	1253	549
1325	446	465	1269	671	696
1156	684	1933	748	1203	2433
1255	1100				

1. Open a New Document and add a Lists & Spreadsheet.

2. Name column A: **intake** and fill the list with the data in the above table.

Doctors involved with this study believe that the subjects had significantly lower than the RDA of calcium. Based on this assumption, we will perform a *t*-test. Don't forget to follow the **4 steps of inference**.

Step 1: State the population and parameter of interest. State the hypotheses.

Population: **Parameter:**

H_0 :

H_a :

[11]Exercise 12.5 "Healthy Bones" – Yates, Moore, Starnes. *The Practice of Statistics, 3rd edition.* W.H. Freeman, 2008, p754.

Step 2: Identify the inference procedure: *One sample t-test* (we don't know σ) and verify the conditions to be able to use it. To use a t-test, we must verify that the data 1) are from a SRS, 2) are Normal, and 3) independent

We have to assume that the data have been collected from an SRS of women ages 18–24 and that the women in the sample are independent (i.e. not related). To check for normality we must either: verify it by the graph (histogram), check that the data follow the 68-95-99.7 rule, or check the normal probability plot.

3. Insert a Data & Statistics page. Create an appropriate histogram.

Describe your graph and comment on the normality of the data.

4. In a separate Data & Statistics page, add the variable **intake** and press (menu), 1: Plot Type, 4: Normal Probability Plot.

Describe the normal probability plot. Do you feel safe stating the data are approximately normal? Why or why not? Support your answer. Provide a rough sketch.

NOTE:

1. If the data have serious deviations from normality, the conclusions from this test may be considered invalid. However, even with a non-normal sample, we proceed with the test. We use extreme caution with making our conclusions about the population based on the results of the calculations.
2. What is *n*? Could we have assumed approximate normality without graphing? If so, why?

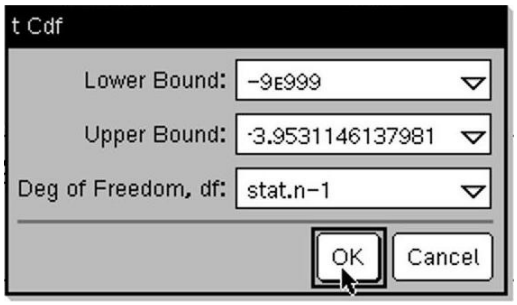
Step 3: Calculate the test statistic and find the p-value.

5. Insert a Calculator page and run the One-Variable Statistics on **intake**.

Record the mean, standard error (s_x), and *n*.

Use these values to find the test statistic, $t = \dfrac{\bar{x} - \mu_0}{\dfrac{s_x}{\sqrt{n}}}$

6. Once you have calculated *t*, on the calculator page press (menu), *Statistics*, *Stat Distributions*, *tcdf*. A dialogue box will appear and you will need to enter the following information: Keep the lower bound as it is, the Upper Bound is your test statistic.



State the p-value.

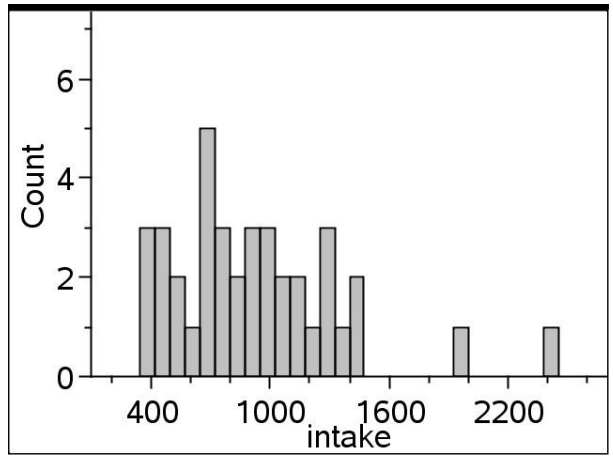
Step 4: Interpretation in context of the problem.

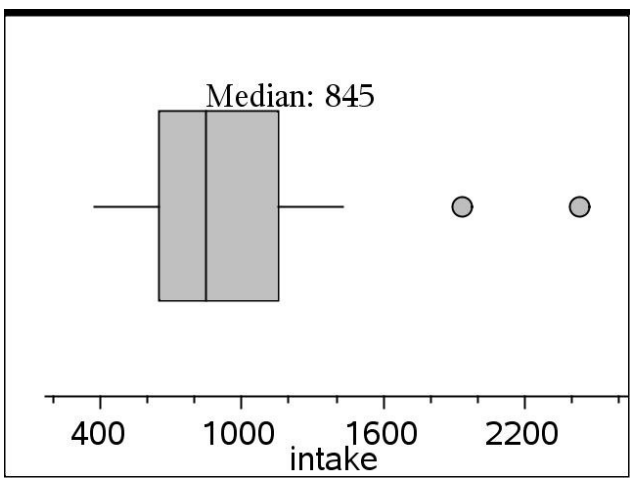
Using the p-value in comparison to an α-level, state your conclusion about the RDA of calcium.

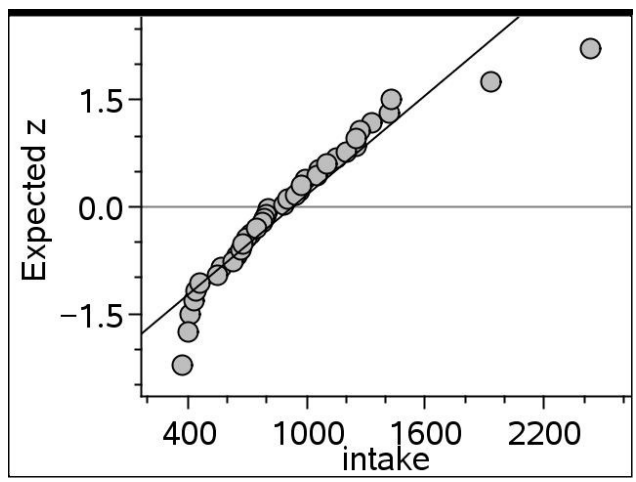
Here are the nspire screens that the student should produce. Using this data, the student should see some deviations from normality. However, given that *n*=38 and that the rest of the data is fairly linear in the normal probability plot, the conclusions can be made.

	A intake	B	C	D
1	808			
2	882			
3	1062			
4	970			
5	909			
6	802			

A1 808







OneVar *intake*,1: *stat.results*

"Title"	"One−Variable Statistics"
"$\bar{x}$"	926.026
"Σx"	35189.
"Σx²"	3.93394E7
"sx := s_{n-1x}"	427.23
"σx := σ_{nx}"	421.571
"n"	38.
"MinX"	374.
"Q_1X"	651.
"MedianX"	845.
"Q_3X"	1156.
"MaxX"	2433.
"SSX := Σ(x−$\bar{x}$)²"	6.75343E6

$\frac{stat.\bar{x}-1200}{\frac{stat.sx}{\sqrt{stat.n}}}$ -3.95311

tCdf(-9.E999,-3.9531146137981,*stat.n*−1) 0.000167

3/99