Let’s pick up where we left off last week. Your PsiPlot spreadsheet should look something like that shown below. You’ll have two columns of data representing Band 16 and Band 20 from Region 1 in the AVIRIS image.

I’ve inserted a Bin column that we will use to specify reflectance ranges to plot in our histogram. The bin values specifically define interval ranges to be displayed in the histogram. When we did this last Tuesday, we used the default bin values.

If you prefer, choose your own bin values. The ones I’ve chosen are rather course, but do illustrate shifts in reflectance.

Bring up the Plot 2D Special Histogram option and generate histograms for the two samples and the specified column for bin values (follow along in class).
Your plots should look something like those shown below.

Visual comparison suggests the reflectance measured by Band 20 may be higher than that measured by Band 16. How would you test whether differences exist? Is you test appropriate?
Using built-in PsiPlot statistical tests.

Follow class demonstration of individual tests. Take notes.

F-Test

<table>
<thead>
<tr>
<th>Column Name</th>
<th>BAND16</th>
<th>Column Name</th>
<th>BAND20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>5007.92857143</td>
<td>Mean</td>
<td>5353.38571429</td>
</tr>
<tr>
<td>Variance</td>
<td>22269.27111269</td>
<td>Variance</td>
<td>36087.3134815</td>
</tr>
</tbody>
</table>

Degrees of Freedom 1: 101
Degrees of Freedom 2: 101
F-Value: 1.62098477
Probability: 0.0012420586

Probability that they're the same, i.e. the alpha level.
t-test

---

T-Test Statistics Report

Data File Name: C:\Classes\GEOG21\Week3\tTest.PAU

FOR COLUMNS BAND16 <--> BAND20

<table>
<thead>
<tr>
<th>Column Name</th>
<th>BAND16</th>
<th>BAND20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>5007.92857143</td>
<td>5135.28571429</td>
</tr>
<tr>
<td>Variance</td>
<td>22843.2711129</td>
<td>36087.3101815</td>
</tr>
</tbody>
</table>

General T-Test Result:
- T-value: -18.28247222
- Probability: 0
- Degree of Freedom: 362

Un-Paired T-Test Result:
- T-value: -18.28247222
- Probability: 0
- Degree of Freedom: 342.76149497

Paired T-Test Result:
- T-value: -98.16697968
- Probability: 0
- Degree of Freedom: 181
- Number of Pairs Used: 182
- Co-Variation: 28167.3520126
- Std Dev: 3.33469710

Probability that they’re the same, i.e. the alpha level.
**Excel Analysis Tool Pak (histograms and statistical tests)**

The following exercise is provided to increase your familiarity with Excel functionality. Excel will generate data histograms similar to those we developed using PsiPlot. Histograms are plotted using Excel’s Analysis Tool Pak. If the Analysis Tool Pak has not been added onto your machine do the following.

First click on tools and examine the drop down window (see below).

Look for Data Analysis. If you can’t find it in the list, you will have to install it. To install the Data Analysis Tool Pak click on Add-Ins and then check the Analysis ToolPak option box (see below). Click OK and the add in should take place.

Once the Analysis tools are installed or if they are already on your machine, proceed to the next page.
Now return to the Tools menu item and you should now see a Data Analysis option in the list. Select the Data Analysis option.

When you click on Data Analysis in the above list, the following window will open up. Highlight the Histogram Analysis Tool and click OK.
The example below uses the pebble mass data discussed earlier, but the procedures are analogous for making histograms of the AVIRIS or other data (your project data, for example). In the following menu, place your cursor in the Input Range box and then with your mouse left-click on the pebble mass data values and slide down selecting the data in that column. When you finish, your screen should look like that shown below.

Click on the BinRange box and select the bin cells provided for you in the worksheet. Finally, click on the Output Range box and then with the mouse sweep through a few rows and columns to give it some cells to output to. Your window should look something like the following.

Also check the Chart Output option box. Then click OK. Your worksheet should look like the following.
Note that the columns of Bin and Frequency appear along with the histogram to the right.

You can edit your histogram plot to increase plot size, change label names, etc.
Left-clicking on the histogram bars will bring up the following window. From there, you can select the options folder and adjust the Gap width in the histogram to eliminate the gaps altogether (set Gap width to 0).

Return to the Tools > Data Analysis > Window

Follow along in class as we run through the F-test.
Brief Notes on Fitting curves to data

In the following example, we assume that the relationship between Age and Depth follows a straight line over the 400 cm to 2500 cm depth range and determine the slope and intercept that define the relationship. The linear equation we wish to fit is \( \text{Age} = \text{Slope} \times \text{Depth} + \text{Intercept} \).

To determine the coefficients of the line (i.e., the slope and intercept) - do the following:
Click on Math- a window drops down.
Click on fitting - another list opens to the side.

Click on Line: \( a \times x + b \); note the similarity of this equation to the one above. Here \( a = \text{Slope} \) above, and \( b = \text{Intercept} \).

The PsiPlot menus are shown at left.

A window opens with a list of column names on the left (the names appearing in the list will vary with your worksheet). In this case you only have two columns of data Depth and Age, and the input data parameters X>> and Y>> default to Depth and Age respectively. The line-fitting window should look like that below. If so - click the OK button.
A panel of statistical information will appear. Note the value of the parameters \( A \) and \( B \).

\( A \), the slope is 1.9 years per cm. This suggests that 2632 cm of sediment will accumulate, approximately, in a 5000 year time period.

Add the regression line to your plot. Remember how? Make sure you highlight the current plot (i.e. click on the background of the plot so that the “handles” appear around its edges) and then click on dataplot to add the best fit line. Note that the variable \( \text{Ind3} \) is our X variable and \( \text{Dep4} \), the Y variable. The Age vs. Depth plot with fitted line is shown below.

Extrapolation of the fitted line over the 0 to 2500 cm depth range corresponds roughly to 4750 years. Note that the fitted line doesn't pass through any of the data points, but is oriented in such a way as to follow the general trend of the data points.

For additional discussion and illustration of regression analysis visit [http://www.geo.wvu.edu/~wilson/qmeth/lect7/regressionlines.pdf](http://www.geo.wvu.edu/~wilson/qmeth/lect7/regressionlines.pdf)
Homework Exercise

You have been given a total of 24 different data sets. Each data set shares some similarity with other data sets either in terms of wavelength (or band number) and surface location. There are six different wavelengths and 4 different surface locations.

- Devise a hypothesis to test certain features of the data. The hypothesis could be posed as a test of differences within a given area between the measurements of reflectance made at different wavelengths. Alternatively a hypothesis could be formed to test for differences between areas measured at the same wavelength.

1. State your hypothesis in writing (one paragraph).

2. State how you will examine your hypothesis. Use the F-test to evaluate the assumption of equality of variances. Use the \( \chi^2 \) to test for normalcy or to compare two distributions. Use the t-test to test for differences in mean reflectance. Compute one or more regression lines to provide another measure of the relationships between reflectance values in different areas (regions/bands, etc. depending on the nature of the hypothesis you wish to test).

3. Present your results and refer to labeled figures.

All reports should be typewritten. The written part of this report (excluding charts and calculations) should be no more than one single-spaced typewritten page (12 font or less). Hand in next lab period and be prepared to present your results in a group discussion.

Due next Thursday (Sept. 18th)