

## SPSS, Importing Data, Regression (Part 2)

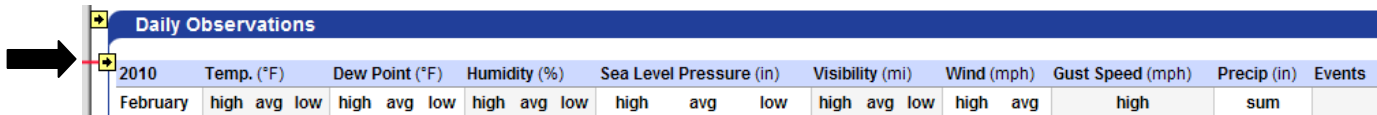
### Import Data

Open a Word document to keep track of your answers for today's work. Be sure to save this Word document on Turing. You do not need to submit this Word document.

► Let's import some data from a table on a web page. The following web page contains weather data for Hillsboro, OR collected in February 2010.

<http://www.wunderground.com/history/airport/KHIO/2010/2/24/MonthlyHistory.html>

Open Excel, Use the **Data** tab, and use the **From Web** button. Put the above URL in the **Address:** bar and press **Go**. You may be prompted to install software, press **Don't Install**. Scroll down to the **Daily Observations** table. Click the yellow box with the black arrow next to **2010** and then press **Import** to bring the data into the current worksheet.



Daily Observations																			
2010	Temp. (°F)			Dew Point (°F)			Humidity (%)			Sea Level Pressure (in)			Visibility (mi)		Wind (mph)		Gust Speed (mph)	Precip (in)	Events
February	high	avg	low	high	avg	low	high	avg	low	high	avg	low	high	avg	high	avg	low	sum	

Since SPSS needs the data in a particular format, we need to edit the spreadsheet. The first row needs to contain the column names. Let's edit row 2 to give meaningful column names then delete row 1.

Let's add a **DayOfWeek** column before the **High Temperature** column. Insert a column, add a heading, and type **Mon** in the second row (Feb 1, 2010 was a Monday). Click and drag **Mon** to fill in the remaining rows. Excel should fill out the days of the week correctly.

Now we need to clean the data. Looking at the spreadsheet, what cleaning of the data might be necessary?

Save this as an Excel file (**Wunderground\_Feb2010**). Open the new Excel file in SPSS. Check the definition of each variable and make any changes you deem necessary.

► Chart 1: Build a chart to show the mean temperature over time, from February 1 to February 28.

What type of chart should you use? Why?

What is on the X-Axis? Why?

What is on the Y-Axis? Why?

Are the values graphed in date order? Why or why not?

Be sure to copy the chart to this Word document.

► Chart 2: Build a chart to show the mean high temperature broken down by Day of the Week, from Sunday to Saturday.

What type of chart should you use? Why?

What is on the X-Axis? Why?

What is on the Y-Axis? Why?

Are the values graphed in the correct day of the week order? Why or why not?

Be sure to copy the chart to this Word document.

What's the problem with the second chart?

What do we need to do to fix this problem?

► Chart 3: Build a chart to show the sum of the precipitation broken down by Day of the Week.

What type of chart should you use? Why?

Be sure to copy the chart to this Word document.

Let's fix your second chart.

Transform | Recode into Different Variables

Select: DayOfWeek      Name: OrderedDayOfWeek

**Change**

**Old and New Values**

Old Value: Sun    New Value: 1      **Add**

**Continue | OK**

**Variable View**

Add values for OrderedDayOfWeek

Rebuild Chart 2 with the days in the correct order.

Further Exercises:

How well does cloud cover predict precipitation? Provide the relevant data and analysis to support your answer.

How well does cloud cover predict mean temperature? Provide the relevant data and analysis to support your answer.

How well does mean humidity predict mean dew point? Provide the relevant data and analysis to support your answer.