# Week 1

Refer to the course map

Welcome to the on line Statistics class! This week we will limit ourselves to a generic introduction to our material, to familiarize with our tools, and to set up ongoing activities that we will work on throughout the Quarter.

The basic information is available here: please, go through all items.

* Syllabus
* Schedule/Map of files and resources
* Tools
* Sources
* Notation (Here is a PDF file you can refer to, if you are having trouble with MathML)

To start our journey, please check **both** the following files, and their links

* A short, simple, slide show (in HTML) about our course
* Quick Introduction (Alternate format: PDF)
* Longer Introduction

During our course, we will use a number of fictitious data files (they have been generated by a computer to present different theoretical possibilities), and will apply statistical tools to this data, as we introduce them. These files come in a variety of formats, to accommodate your needs. You can download them now, and be ready to work on them.

## Objectives for this week

* Familiarize yourself with the course and its layout:
  + Check how tonavigate the Map course page, the weekly assignment pages bck and forth (do not forget the < BACK > button in your browser)
  + Similarly, make sure there are no problems navigating other pages, as the come up in the links
* Familiarize yourself with the tools: be able to do basic manipulation of the tools we will use.
  + **Specifically, acquire and try out your spreadsheet: depending on your choice, you will have different ways to work on the data. Please, explore your program, check out its instructions, and write me with specific questions**
  + In particular, explore actions like the following:
    - Add sequences of numbers
    - Add the squares of the same sequences of numbers
    - Do arithmetic operations on the results
    - Apply functions that your spreadsheet provides to those numbers, and further manipulate the results Specific functions that we will end up using include, among others, normdist, normsdist, norminv, normsinv, tdist, tinv. We will be explaining them as we introduce their definitions
    - *Gnumeric* would do most of the hard work for you, but it is not a bad idea to set up the needed calculations yourself - which you will have to do anyway, if you use another spreadsheet
* Explore the Internet for a data set concerning a topic or issue that is of interest to you, try to identify one, and discuss your choice with your instructor for approval.

You might want to enroll in a group, as suggested in the Maps page, and below, on this page. This might be a way to make it easier to collaborate or, at the very least, share suggestions, tips, experiences, and so on. An initial list of groups is listed there, and has been activated on Blackboard, but suggetsions for adding, changing, specifying, and so on, are welcome.

## Reading Material

Please, go through the informational files about the course. Here is a reminder of the list of informational files:

* Syllabus
* Schedule/Map of files and resources
* Tools
* Sources
* Notation (Here is a PDF file you can refer to, if you are having trouble with MathML)

If you have any questions, write me. Also, take some time to make sure you are familiar with the basic operations of your spreadsheet: going through the documentation (Help files, Manual), as well as looking for tutorials, How-Tos, and similar sources on the web, will make it much easier for you.

## Assignment for the week

There is no written assignment to turn in this week. To familiarize yourself with your spreadsheet, you should try, at least, the following:

* Set up a column of numbers (for example, as a first application of functions, set each cell equal to =rand() (which will produce a "random" number between 0 and 1), or, if you prefer integers, =randbetween(a,b), where *a* and *b* are any two numbers (this will produce an integer greater than *a*, and less than *b*)
* Have you spreadsheet compute the sum of these numbers
* Compute the square of each number, and add all the squares
* Go to the function list, and choose one or, preferably, more, and apply it to the list or to each number, depending on the function. For example, you could select the functions that fall under the "statistics" header, such as average(),stdev(), var(), max(), min() and so on (it does not matter whether you know what they do or not)

Also, work out the exercises you will come across in the course material you are going through this week. Don't miss the conclusion of the Notation file (Here is a PDF version of the file), where a very useful formula is presented (it will turn handy in "pencil and paper" work, later). Ask any question that may come up!

## Start On Your Project

As an experiment, you are urged to choose one or more data sets from publicly available sources (presumably, via the Internet), and analyze them with our tools, as we introduce them. It would be very good if you could work as teams of 2 - 3 members, even if an on line class does not make this very easy. In any case, the idea is for you to choose data relevant to a topic that is of interest to you.

As a generic help, we may divide you in groups (as provided by Blackboard), according to the general area of interest you choose. Even if you should not team up, exchanging info and thoughts with people addressing similar issues could be useful. As a start (further suggestions for additions and/or refinements are welcome), we could consider groups looking for data in

* Economy and Finance
* Demographic Data
* Environmental Sciences
* Health Sciences
* Sports

One possibility (but feel free to work as you see best fit) is to start by searching the web with your favorite search engine for data on a topic that interests you. As a simple example, a Google search for "sports statistics" will lead to a results page, where, for example, you will find a link to [this page, organized by the American Statistical Society,](http://www.amstat.org/sections/sis/sports%20data%20resources/) with further links to a variety of resources with statistical data on many sports.

Please, submit your choice, so we can discuss whether it is appropriate and feasible.