PHPM 672 Assignment #1: Hello World!
Due date: Submit in E-Campus by 11:59pm Monday 1/22

Submission. Submit on E-campus by 11:59pm the day before the class they are due.
- Commented code (SAS: lnameN.sas, where N indicates the assignment number and lname is your last name)
- Output from your code (SAS: lnameN.log & lnameN.lst or lnameN.html)
- Readme.txt file answers to questions

Homework Assignments. There will be a homework assignment due approximately every other week. Homework assignments and related materials will be made available on the course website. **I strongly advise students to begin to work on their homework assignments soon after they are assigned, so that if they need help from the instructor they have time to ask and receive assistance.** There will be *six (6)* regular HW assignments during the semester. The overall weight of the HW assignments on the computation of the final course grade is *48%, HALF the grade.*

All assignments are due at 11:59pm the day before the class they are due.

Late Assignments. Each student will be allowed one late assignment, due 7 days from the due date. NO other late assignments or make up will be accepted.

Assignments and Exam Schedule (Minor modifications may be made as needed with notice)

	Date	Assignment Given	Assignment Due
1	1/18 (THUR)	Assignment 1	
2	1/23	Assignment 2	Assignment 1
3	1/30		
4	2/6	Assignment 3	Assignment 2
5	2/13		
6	2/20	Assignment 4	Assignment 3
7	2/27		
8	3/6	Midterm	
	3/13	SPRING BREAK	
9	3/20	Assignment 5	Assignment 4
10	3/27	Assignment 6	Assignment 5
11	4/3		
12	4/10	Final Project (FP)	Assignment 6
12	4/12 (THUR)		FP MileStone 1
13	4/17		FP MileStone 2
14	4/26 (THUR)	LAST DAY OF CLASS*	Final Project

* May 1 seems to be a redefined class to a Friday class.

Collaboration: Collaboration on assignments, in class labs and homework, IS encouraged. However, what you hand in must be in your own writing/typing. Good scholarship requires that all collaborations must be acknowledged. Thus, if you collaborate on the solution of the problem set, we expect that you list your collaborators at the top of the page. Collaboration on in-class evaluations (quizzes, mid-terms, and the final exams) is, of course, a violation. This includes a discussion of questions on a quiz, midterm, or final with students from sections that has not yet taken the evaluation.

Plagiarism: If you consult any outside sources when doing your work, you are expect to further document these sources. Give credit where credit is due. Plagiarism will not be tolerated.

All handed in homework should state at the top any assistance with debugging and programming, as well as citations of any program segments copied from a website.

Seeking Help: Language classes are cumulative; don't fall behind. Ask for assistance from your fellow students or from the instructor, especially if you find yourself struggling. But remember in the end, programming classes are like skill building math classes. Either you know how to do them or not. No amount of watching others or the instructor program or debug it will suffice for you to build these skills. Only your hard work to work through them will build your skills. It is a lot of work to build these skills, but data skills are highly valuable in the job market in the modern digital world, so it will be well worth your effort. Furthermore, once you learn to think in this manner, it's not something you forget.

Elegance: There is always more than one way to say something, but some ways will be more "elegant" than others. This is true for computer languages as well. You will learn to recognize elegant expressions as you become more familiar with a language and use the elegant idioms as you become more skilled.

Recommended readings for this assignment
In the following sections, if I copied and pasted sections from any of these websites, it is italicized.
1. http://www.ats.ucla.edu/stat/sas/

Assignment 1: Hello World!

In this assignment, you will write your first SAS code. It is a very simple assignment, so you should be able to complete this in the computer lab if you do not have access to the software elsewhere.

Decide your work environment for the class (MOST IMPORTANT)
1. You need to first decide your mode of work for this class. The overhead of working "temporarily" is too high for this class. So you need to decide which computer you will be using for this class for the full semester. The following is a list of requirements for the computer you choose to use.
    a. It must had SAS installed. All students can install SAS on your personal computers for free, using the SAS University Edition (https://www.sas.com/en_us/software/university-edition.html). But sometimes you will need help getting that installed. So if this is what you plan to do, make sure to get SAS installed ASAP. If you are using computer owned by the university (e.g., your desktop in the office) you will have to figure out who will pay for the license. There is no support for this.
    b. The VOAL is another alternative. Using VOAL you are able to access SAS using your personal or university computer as long as you have internet access.
   c. It is best if there is hard disk space for you to store your work for the whole semester. If not, such as using the lab computers, you will have to purchase a USB external harddrive to store class material
   d. You need to be able to access it during class. That is, it is a lab computer, or you have your laptop with you, or you remote desktop into the computer from the lab.
   e. If you do not have better options, you will have to use the lab computers. This means you will buy the USB to store class material.
2. If at all possible, we recommend that you decide and setup, then use that computing environment to do this assignment. This will make it possible for you to focus on learning the rest of the semester.

Setting Up
3. Locate the software (SAS) on the computer and execute it
4. Create a working directory, often referred to as pwd (present working directory), where you will work on this assignment. For example, "assign1".
5. When you are writing programs, it is often a good idea to put your code in a different directory from your datasets, because if you are writing a lot of code, or working with a lot of datasets, you will stay more organized. So, create a subdirectory called "data" in the pwd

6. Now, get a good text editor. For example, portable apps notepad++ or notepad2 are mentioned below. You should be able to unpack this software into your pwd or USB and run from there by double clicking on "Notepad2…" if you cannot install on the lab computer.
7. http://portableapps.com/apps/development/notepadpp_portable
   http://portableapps.com/apps/development/notepad2-mod_portable
8. Now you are READY!

Getting Data: Download Flu Trends Data from Google Flu Trends
1. Go to http://www.google.org/flutrends
2. Choose "United States" from dropdown on the left as "national"
3. Click "download data" link on the left (leave the second drop down)
4. Save as "flu.csv" in your data directory
5. Open this csv file using the text editor
6. Cut and Paste the comments on top (11 lines) into a different text file ("readme.txt"). This is documentation on what the data is for your reference later, in case you have to come back to this. The csv file should start with "Date, United States …"
7. Don't forget to save your files.

What you will do with the data via programming:
1. Import the data into the software (see below)
2. Find out how many variables (columns) and observations (rows) are in the data, and what is in the data. Note you have to get the software to output this, even if you know the answer to this question.
3. Print the first 10 observations
4. [EXTRA CREDIT (+1)] If you feel really ambitious, try to find the mean of each state. If you can't figure out how to do this after trying a bit, don't worry. We will learn this later in class.

P1. SAS programing:

1. Open a new file in the text editor. In this file, type:
   ```
   *This program imports, describes, and displays the dataset;

   * EXAMPLE: libname data 'C:\phpm672\assign1';
   libname data 'LOCATION OF FOLDER'; *you have to type in location;

   *you have to type in location and filename;
   * EXAMPLE: datafile= C:\phpm672\assign1\flu.csv;
   proc import out= data.flu
               datafile= "LOCATION AND NAME OF FILE"
               dbms=csv replace;
        getnames=yes;
        datarow=2;

   proc contents data=flu;
   proc print data=flu(obs=10);
   ```

   Save your program as "flu.sas" and exit.

2. Let's look at this program line by line.

   The first line of the program is a comment. Comments begin with a "*", and end with a ";" You can type anything you want in a comment, but generally comments give information about your program.

The next line (libname) tells SAS where to find the dataset you will be using, and call that folder *data*.

The next five lines import the csv dataset, and creates a SAS dataset called flu in the folder pointed to by *data*, which contains the data from flu.csv.

The line after that (contents) gives a description of the data.

The final line prints the first 10 observations in the data.

3. Use SAS to run your program. Note that the command you use to run SAS depends on the system you are using. If you double click on the sas program it should run sas.

4. Now, save out the log and lst.

P2. Try using point and click
Now, try doing the same task you did with point and click mode (without writing the program) in the SAS software.

P3. Questions:
Answer these questions in a text file called readme.txt (convention for documentation of programs).
1. State the computing environment you used to this assignment.
2. Describe what data (flu.csv) you have in no more than three sentences
3. Write one thing you liked better and one thing you did not like about using programs compared to point and click.
4. Working environment: It is important for you know the set up your work environment to do good programming. Programming is very difficult and time consuming if you don't have your work environment set up. By now, you should have figured out what environment you are going to program for this class. Answer the following questions, so I know you are set up. If you have more than one mode of working, answer the questions based on your primary mode
    a. What computer are you working on? (e.g., home desktop, personal laptop, lab computer, virtual lab etc)
    b. What version of SAS is installed on the computer. Include OS.
    c. If this is not a lab computer, how are you accessing the computer during class? (i.e. remote access via remote desktop)
    d. Where is your working directory on the computer? Type the full path to the root folder for this classwork, and then list the current file structure (list all folders) of that folder.

Below is an example answer for myself,
    a. Lab comptuer
    b. windows base SAS v9
    c. NA
    d. Z:\phpm672 (this is a SPH network drive accessible from the lab computers available to faculty. Z drive is mapped to my personal home folder)
        • Z:\phpm672\assign1
        • Z:\phpm672\assign2
        • Z:\phpm672\assign2\data