Math 585, SAS on Windows Handout: SAS is a statistical software package that may be used in this course. You may want to use a flash drive. To use SAS on windows (PC), use the following steps. SAS may be on SAS is on computers 11-25.

1) Google David Olive, get on his website, click on the M585 website, scroll to th bottom click on the mult link, to get to (http://parker.ad.siu.edu/Olive/mrsashw.txt). Highlight the SAS program. Hit Ctrl–c. Type sas in the search window near the lower left icon. Double click the SAS icon. The editor window is the lower window. Click on that window, then hit Ctrl–v to paste in the program. Then run>submit. Output will appear in a few minutes. If your computer does not have SAS, go to another computer.

2) A window with a split screen will open. The top screen says *Log-(Untitled)* while the bottom screen says *Program Editor-(Untitled).*

When you are done entering the program as descibed in 3) below, you may want to save the program as hw5df.sas on your flash drive (J: drive, say). (On the top menu of the editor, use the commands "File > Save as". A window will appear. Use the upper right arrow to locate "Removable Disk (J:)" and then type the file name in the bottom box. Click on OK.)

3) Copy and Paste the the following SAS program from the bottom of (http://lagrange.math.siu.edu/Olive/mrsashw.txt) into the SAS editor window. The *ls* stands for linesize so *l* is a lowercase *L*, not the number one.

```
*SAS Institute (1985, p. 626-630);
options ls = 70;
data Temperature;
      title 'Mean Temperature in January and July for Selected Cities ';
      input City $1-15 January July;
      datalines;
Mobile         51.2 81.6
* more data;
   ;

 title 'Mean Temperature in January and July for Selected Cities';
   title2 'Plot of Raw Data';
   proc plot;
      plot July*January=City/vpos=36;
   run;

   title 'Mean Temperature in January and July for Selected Cities';
*leave cov out to use correlation matrix;
   proc princomp data=Temperature cov out=prin;
      var July January;
   run;
   proc plot;
     plot prin2*prin1/vpos=26;
   run;
```

(If the file is saved on Removable Disk (J:), get back into SAS, and from the top menu, use the "File> Open" command. A window will open. Use the arrow in the NE corner of the window to navigate to "Removable Disk (J:)". (As you click on the arrow, you should see My Documents, C: etc, then Removable Disk (J:).) Double click on **hw5df.sas**.)

4) To execute the program, use the top menu commands "Run>Submit". An output window will appear if successful.

If you were not successful, look at the *log window* for hints on errors. A single typo can cause failure. Copy and paste the SAS program in *Word* or *Notepad* and make corrections. Occasionally you can not find your error. Then find your instructor or wait a few hours and reenter the program. Click on *log* window, then click on Edit and select Clear All. Use similar commands to clear the *Editor* and *Output* windows.

5) To copy and paste relevant output into *Word,* or *Notepad*, click on the output window and use the top menu commands "Edit>Select All" and then the menu commands "Edit>Copy". In *Notepad* use the commands "Edit>Paste". Then use the mouse to highlight the relevant output. Then use the commands "Edit>Copy".

Finally, in *Word*, use the commands "Edit>Paste".

6) This point explains the SAS commands. The semicolon ";" is used to end SAS commands and the "options ls = 70;" command makes the output readable. (An "*" can be used to insert comments into the SAS program. Try putting an * before the options command and see what it does to the output.) The next step is to get the data into SAS. The command "data Temperature;" gives the name "Temperature;" to the data set. The command "input City $1-15 January July;" says the first variable City is a text variable entered on the first 15 spaces of each line, and the 2nd variable is January and the third variable is July. The data gives average January and July temperatures for 64 cities. The command "cards;" or "datalines;" means that the data is entered below. Then the data in entered and the isolated semicolon indicates that the last case has been entered. The commands "proc plot; plot July*January=City/vpos=36;" make a scatterplot of the July vs January temperatures using symbols related to city. The commands "proc princomp data=Temperature cov out=prin; var July January;" does principal components using the covariance matrix with the two given variables. Output is stored under the name 'prin'. Use "proc princomp data=Temperature out=prin; var July January; " with the "cov" omitted to do principal components with the correlation matrix (the default). The commands "proc plot; plot prin2*prin1/vpos=26;" make a scatterplot of the first versus second principal component.

SAS is the industry standard for statistical packages, and the following references may be of interest (google SAS User's Guide: Statistics).

Cody, R. (2007), *Learning SAS by Example: a Programmer's Guide*, SAS Institute, Cary, NC.

Cody, R.P., and Smith, J.K. (2006), *Applied Statistics and the SAS Programming Language,* 5th ed., Pearson Prentice Hall, Upper Saddle River, NJ.

SAS Institute (1985), *SAS User's Guide: Statistics*, Version 5 , SAS Institute, Cary, NC.