1. Consider the following SAS program, available in the file `assign3_prob1.txt` downloadable from the Stat479 course page Homework Assignments link. This program inputs information about hospitals in cities in different states, such as the size of the state and the cities (i.e., population), median income and median housing costs, number of admissions and the number of beds, using three different data types 1, 2, and 3. Answer the questions below:

```sas
data hospital;
retain Form_Id State Stpop CitySize Income Housing Admit Beds;
input Form_Id $3. @5 Type $1. @ ;
if Type =’1’ then input @8 State $2. Stpop 10.;
if Type =’2’ then do;
  input @8 CityPop 8. Income 5. Housing 6.;
  if CityPop<60000 then CitySize=’Small’;
  else CitySize=’Large’;
end;
if Type =’3’ then do;
  input @8 Admit 6. Beds 4.;
  output;
end;
drop Type CityPop;
datalines;
   v12 1 IA 1708232
   v12 2 53620 5240 14236
   v12 2 5126 178
   v12 3 3364 134
   v12 3 4857 184
   v12 1 KS 1575899
   v12 2 86610 4879 18154
   v12 3 3916 156
   v12 3 5527 182
   v12 3 12139 351
   v12 3 8257 238
   v12 2 36574 3754 12739
   v12 3 3465 112
   v12 3 4576 142
;
proc print;
run;
proc means data=hospital noprint;
class State CitySize;
var Admit Beds;
output out=stats mean= Av_Adms Av_Beds std=S_Adms S_Beds;
run;
proc print data=stats;
run;
```

(a) Show the contents of the PDV immediately after processing the first line of data.
(b) Show the contents of the PDV immediately after processing the second line of data.
(c) Show the contents of the PDV immediately after processing the third line of data.
(d) Show the contents of the PDV immediately after processing the fifth line of data (before the observation is output to the SAS data set).
(e) Display the first observation written to the SAS data set.
(f) Run this program and turn-in the output only.
(g) Examine the output produced by the second proc print step. Describe the statistics printed in each line of this output, i.e., explain what the computed numbers are in each line. Make sure that for each value of the _TYPE_ variable, that you identify the group of observations used to compute the statistics that appear in that line.

2. This problem uses the demographic data set on countries given in Table A.4 of Appendix A (available for downloading under the name demogr.txt) from the homework assignments page. Copy this to your folder in the U drive as a text file. To access this data in your SAS program, include an appropriate infile statement. In your data step, use the following input statement to read these data:

```
input Country $20. Birthrat Deathrat Infmort Lifeexp Popurban Percgnp Levtech Civillib;
```

(a) Input the data and create a SAS data set named world (make sure that this dataset is saved permanently in your folder). Use SAS statements in the data step to create two category variables to be added to the data set as described below:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Groupings</th>
<th>Category Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infmort</td>
<td>&lt; 24 = 1</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>24−73 = 2</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>≥ 74 = 3</td>
<td>High</td>
</tr>
<tr>
<td>Levtech</td>
<td>&lt; 24 = 1</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>≥ 24 = 2</td>
<td>High</td>
</tr>
</tbody>
</table>

Provide more descriptive labeling for all variables using a label statement. You may use proc print to check whether your SAS program works; however, it is not required to turn-in this output. Just turn-in a Word output of your SAS program.

(b) Those countries with Birthrat exceeding 40 and Percgnp less than 1000 were selected for an advanced study. Add a proc print step to the SAS program in Part (a) to obtain a SAS listing (i.e., a SAS report) that contains only the variables Deathrat, Infmort, Lifeexp, and Popurban for these countries. You must not create any additional SAS data sets to do this and make sure that all variables are labeled and all observations identified by the country name in the report.

(c) Write a new SAS program to access the SAS data world from your folder in a proc means step. This step must contain appropriate var, class, and output statements, required to create a SAS data set named stats1. Suppress any printed output from this proc means step. The data set stats1 must contain sample means and standard errors of the means of the variables Birthrat and Deathrat calculated separately for each of the six groups defined by combinations of levels of the category variables Infgrp and Techgrp. Use the types statement to ensure that statistics are calculated only for combinations of levels of Infgrp and Techgrp. Obtain a listing of the data set stats1 using a proc print step with labels for all variables appearing in this listing.

NOTE: Insert ODS statements to direct the output from your proc steps to an appropriately named rtf file in your U drive Stat479 folder. Add a listing of each of your program to each of these rtf files, label appropriately, and turn-in a printed versions of the rtf files as your solution to this problem (this should be only about two printed pages). Make sure that you do not delete the SAS dataset world from your folder as it may be needed for doing future assignments/quizzes.

**Due Tuesday 29, September 2015 (turn in at the beginning of the class)**