Instruction

- Please save your solution in **one single file** with the *.sas* extension.

- Please e-mail me (arthur.x.li@me.com) your assignment by the due date. You only need to send me your SAS code.

- Please do not send me the output and the log files.

- Please do not send me a zip file.

- When writing your homework, please follow the template on the last page of the assignment.

Problems from Chapter 5 & Chapter 6

This homework will be based on the materials from Chapters 5 & Chapter 6 of the textbook. I strongly recommend that you do all the exercises at the end of each chapter. However, you are only required to turn in one problem of your choice from each chapter. Since you can obtain the solution from the publisher's website, these two problems will be graded on completion only.

Problem 1

You are given the SAS data set SPEED (*speed.sas*), created by running the program below. Create a new data set SPEED2 from SPEED, with some new variables. The new variables LX1 - LX5 are the natural log of the variable X1 - X5, and variables SY1 - SY3 are the square roots of the variables Y1 - Y3. Use arrays to create the new variables.

```
data speed;
    input X1-X5 Y1-Y3;
datalines;
1 2 3 4 5 6 7 8
11 22 33 44 55 66 77 88
;
```

Problem 2

You will work with the *dna.sas* file that reads 15 DNA sequences (See below). The length of each sequence is 60 characters. Based on these DNA sequences, create 60 variables, D1 - D60. D1 will hold the DNA at the first position, D2 will hold the DNA at the second position, and so on. You must use array processing to complete this problem. Hint: use the SUBSTR function.

```
data dna;
    length dna $ 60;
    input dna $;
datalines;
TGGAAGGGCTAATTTGGTCCCAAAAAAAGACAAGAGATCCTTGATCTGTGGATCTACCACA
TGATTGGCAGAACTACACACCAGGGCCAGGGATCAGATATCCACTGACCTTTGGATGGTG
CTTCAAGTTAGTACCAGTTGAACCAGAGCAAGTAGAAGAGGCCAAATAAGGAGAGAAGAA
CAGCTTGTTACACCCTATGAGCCAGCATGGGATGGAGGACCCGGAGGGAGAAGTATTAGT
```

```
GTGGAAGTTTGACAGCCTCCTAGCATTTCGTCACATGGCCCGAGAGCTGCATCCGGAGTA
CTACAAAGACTGCTGACATCGAGCTTTCTACAAGGGACTTTCCGCTGGGGACTTTCCAGG
GAGGTGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATGCTACATATAAGCAGC
TGCTTTTTGCCTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTG
GCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTCAAAGTAG
TGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAG
TGTGGAAAATCTCTAGCAGTGGCGCCCGAACAGGGACTTGAAAGCGAAAGTAAAGCCAGA
GGAGATCTCTCGACGCAGGACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCG
GCGACTGGTGAGTACGCCAAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGTG
CGAGAGCGTCGGTATTAAGCGGGGGAGAATTAGATAAATGGGAAAAAATTCGGTTAAGGC
CAGGGGGAAAGAAACAATATAAACTAAAACATATAGTATGGGCAAGCAGGGAGCTAGAAC
;
```