CSCI 1530 Computer Principles and Java Programming 2014-2015 Second Term
Department of Computer Science and Engineering
The Chinese University of Hong Kong

Due date: 21 April 2015 (Tue) **Assignment 6** Full mark: 100
Expected normal time spent: 4 hours

Statistics on Array Data and File I/O

Aims: 1. build a practical standalone analytic tool using Java;
2. practise using looping constructs on an array;
3. work with file I/O and exception handling

Task: Finding some statistics of Hang Seng Index (HSI) daily close.

Background:

Hang Seng Index (HSI) is a key economic and financial indicator in HK. It is also well-known and widely referenced globally. We are going to do some statistics on the historical records of the HSI Daily Close figures[1]. In this assignment, we assume there are at most 4000 records to process. We shall read data from an online source into some array(s). Then, find some statistics accordingly.

Requirements:

The program should read and process raw HSI trading data (in chronological order) from this URL:

http://www.cse.cuhk.edu.hk/csci1530/assignment/HSI.txt

Note that the trading days may not be consecutive calendar dates. This is not an issue because the program shall count number of trading days (e.g. N) according to the data available in the series. Thus the ascending time stamps (dates formatted in YYYYMMDD) in the raw data may be considered as tags and keys, for identification purpose only. The number of rows in the data set really matters, but not the number of actual calendar days spanned.

Firstly, the program asks the user for a single input value which is 8-digit integer in YYYYMMDD:

```
Start Date
```

The user input is assumed to be valid, i.e., it must fall within the available range in the data set and the start date is a trading day with proper data. Thus, no input validation is required. The program then calculates the max, the min, the average and the median of the HSI Closes from the Start Date till the end of the available data set.

Sample run: (user input is underlined in RED)

```
Start Date: 20001215
HSI from 20001215 to 20150409
Number of records: 3572
Max: 31638.22
Min: 8409.01
Average: 17924.20285
Median: 19050.45
```

If the number of records processed is odd, the median should be the middle value. In another case, if the number of records is even, take the mean of the middle-two values.

[1] Data are primarily retrieved from TR4DER - Hang Seng Index - Hong Kong [^HSI] Historical Prices on 9 Apr 2015 http://www.tr4der.com/historical-prices/%5EHSI/

Exception handling: in case of any exceptions and/or unexpected conditions, the program should simply **terminate at that point**.

Procedure:

1. Create a new project **HSI**, with a new class **HSI** under package name **tool**. It may import/ make use of the many Java API packages and classes for file I/O or web data access. NetBeans can help to "Fix imports" for us.

2. You should use the **Scanner** class to obtain user inputs from the console **System.in** interactively. The program need not validate (check) the user inputs.

3. Raw data from the given URL is contained in a plain text file. The first two lines are textual headers and may be skipped and ignored safely. Subsequent lines are tab-separated data containing an integer date stamp and a double-type number on each line. We assume the data set is well-formatted. Sample code for data reading (in your work, you shall use some array(s)):

```
public static void main(String[] args) throws Exception {
   Scanner userInput = new Scanner(System.in);
   int startDate = userInput.nextInt();

   String addr;
   addr = "http://www.cse.cuhk.edu.hk/csci1530/assignment/HSI.txt";
   URL link = new URL(addr);
   Scanner dataStream = new Scanner( link.openStream() );
   String  line1 = dataStream.nextLine();
   String  line2 = dataStream.nextLine();
   int      day1 = dataStream.nextInt();
   double   hsi1 = dataStream.nextDouble();
   if ( dataStream.hasNextInt() ) {       // check for end-of-data
     int    day2 = dataStream.nextInt();
     double hsi2 = dataStream.nextDouble();
     System.out.println("Demo:- Close on " + day2 + ": " + hsi2);
     System.out.println("Demo:- read just 2 records as a demo");
   }
}
```

4. You are strongly advised to open and examine the data file using a browser, your favourite text editor as well as Microsoft Excel (save the file and drag-and-drop it to Excel.) You may also validate your program's output using Excel some formula.

5. We just print the statistics using Java **System.out.println(double)**, without concerning the number format in the output.

6. Hint: you may read dates and HSI closes into two arrays. When the user specifies a Start Date, search for the specified date in the dates array and keep the index (row number.) Then find the statistics in the HSI close figures array. You may need to sort the array for finding the statistics. Alternatively, you may define a class Record and create an array of Record objects to proceed.

7. Thoroughly Test Run your program with different data sets (you may create your own local file and/or URL for testing purpose during the development) and user input values. **Be reminded to apply the default official URL before you submit your finalized work.**

Your Tasks:

1. Include the following header comment block in your Java source file **HSI.java**:

```
/**
 * CSCI1530 Assignment HSI
 *
 * I declare that the assignment here submitted is original
 * except for source material explicitly acknowledged,
 * and that the same or closely related material has not been
 * previously submitted for another course.
 * I also acknowledge that I am aware of University policy and
 * regulations on honesty in academic work, and of the disciplinary
 * guidelines and procedures applicable to breaches of such
 * policy and regulations, as contained in the website.
 *
 * University Guideline on Academic Honesty:
 *   http://www.cuhk.edu.hk/policy/academichonesty/
 * Faculty of Engineering Guidelines to Academic Honesty:
 *   http://www.erg.cuhk.edu.hk/erg-intra/upload/documents/ENGG_Discipline.pdf
 *
 * Student Name: xxx <fill in yourself>
 * Student ID  : xxx <fill in yourself>
 * Date        : xxx <fill in yourself>
 */
```

2. ZIP the project folder **HSI** and Submit the file **HSI.zip** via our Online Assignment Collection Box on Blackboard <**https://elearn.cuhk.edu.hk**>

Marking Scheme and Notes:

1. The submitted program should be free of any typing mistakes, compilation errors and warnings. Comment/remark, indentation, style are under assessment in every programming assignments unless specified otherwise. Variable naming, proper indentation for code blocks and adequate comments are important.

2. Remember to do your submission before 6:00 p.m. of the due date. No late submission would be accepted.

3. If you submit multiple times, **ONLY** the content and time-stamp of the **latest** one would be counted. You may delete (i.e. take back) your attached file and re-submit. We ONLY take into account the last submission.

University Guideline for Plagiarism

Attention is drawn to University policy and regulations on honesty in academic work, and to the disciplinary guidelines and procedures applicable to breaches of such policy and regulations. Details may be found at http://www.cuhk.edu.hk/policy/academichonesty/. With each assignment, students are required to submit a statement that they are aware of these policies, regulations, guidelines and procedures.

Faculty of Engineering Guidelines to Academic Honesty

MUST read: http://www.erg.cuhk.edu.hk/erg-intra/upload/documents/ENGG_Discipline.pdf (VPN required if you are not in the CUHK campus network)