Comp 333   Project #1 (FORTRAN to Java Conversion Project)

Due in class: Thursday Feb 5, 2015
Points: 20

You must work with your assigned team member. This is the team member you worked with in class on Thursday, Jan 22. Each group should submit only one project with both team member names on the title page of the project.   All groups are teams of two, with one exception. If you do not have a team, let your instructor know right away.

1. Study the attached Linear Regression Fortran program. This program finds the line y = ax + b that best fits the scatter plot of a set of data pairs (x_i, y_i). It also finds the correlation coefficient. Using the programs variables, S is the slope, B is the y-intercept and R is the correlation coefficient.

2. Translate the FORTRAN program (more or less line for line, including the format statements and comments) to a Java program.

    a. **Do not use gotos. Do not use subprograms.** Use for loops, while loops and if statements to replace the gotos and do loops. Do not create your own version of a linear regression algorithm.
    b. You must compute and print all of the values and strings printed in the program using the same formats. You will need to figure out what the FORTRAN formats mean. Check on the internet.
    c. Translate DIMENSION LITERL(2) into a java statement that creates an array of 2 characters.
    d. FORTRAN uses arrays that start with index 1. Change the program in Java so that the arrays start with index 0.

3. Test Case #1: Run the following test case: N = 7 with pairs   -8,-1 ; 1,4; 0.5,-4; -2,1; 4,2; 10,10; -10, -2.3 . Run the test case, taking the option of reviewing the data points. **Numeric answers: S = 0.514368 ;   B = 1.716379; R = 0.758499**. On another sheet of paper, plot the 7 points and plot the best fit line. Note that generally most of the data points will not fall onto the best fit line.

4. Test Case #2: N= 3 with pairs 0,1: 1,-3; 2,-7. Plot data points and best fit line. Why do the data points fall on the best fit line in this case?

5. Test Case #3: N = 10 with pairs : -4,5; -3,3; -2,1; -1,-1;   1,-5; 2,-7; 3,-9; 4,-11; -1.5,0; 1.5,-6.   Plot data points and best fit line.

6. **Hand in in class on Feb 5: (stapled together)**
   a. Title Page with course, project, date and team member names
   b. Java source code with the printout of Test Case #1 included at bottom of .java file
   c. Results of Test Case #1 (program printouts and plot)
   d. Results of Test Case #2 (program printouts and plot)
   e. Results of Test Case #3 (program printouts and plot)

```
      The Steve Sharpe (University of Bridgeport)
C           Irvin Levy (Gordon College) Linear Regression Package Modified
C
C           *  *  *  *   VARIABLE LIST *  *  *  *
C                    S1 - Sum of X values
C                    S2 - Sum of Y values
C                    S3 - Sum of X * Y
C                    S4 - Sum of X * X
C                    S5 - Sum of Y * Y
C                    S  - Slope
C                    B  - Y-intercept
C                    R  - Correlation coefficient
C
C
C
C
          DIMENSION X(50), Y(2,50), LITERL(2)
          DOUBLE PRECISION S1,S2,S3,S4,S5,T,S,B, R,BBAR
          WRITE (5,10)
 10       FORMAT('0',1X,'*   *   *  LINEAR REGRESSION ANALYSIS  *   *   *',//)
          WRITE (5,20)
 20       FORMAT(1X,'HOW MANY PAIRS TO BE ANALYZED?'$)
          READ (5,*) N
          IF (N.GT.50) GOTO 70
          WRITE (5,30)
 30       FORMAT(//1X,'Enter one pair at a time')
          WRITE (5,40)
 40       FORMAT(1X,'and separate X from Y with a comma .'//)
          WRITE (5,50)
 50       FORMAT(1X,'Enter pair number one : '$)
          READ (5,*) X(1), Y(1,1)
                  DO 60 I=2,N
                  WRITE (5,55) I
 55               FORMAT(1X,'Enter pair number',I3,' : '$)
                  READ (5,*) X(I), Y(1,I)
 60               CONTINUE
          GOTO 90
 70       WRITE (5,80)
 80       FORMAT(1X,'At present this program can only handle 50 data pairs.')
          STOP

 90       WRITE (5,100)
 100      FORMAT(1X,'Would you like to print the data? '$)
          READ (5,110) LITERL(1)
 110      FORMAT(A1)
          IF (LITERL(1).EQ.'N') GOTO 200
                  DO 130 I=1,N
                  WRITE (5,120) I, X(I), Y(1,I)
 120              FORMAT(1X,'DATA PAIR',I3,').  ',2F20.10)
 130              CONTINUE
```

```
200              DO 210 I=1,N
                 S1=S1+X(I)
                 S2=S2+Y(1,I)
                 S3=S3+X(I)*Y(1,I)
                 S4=S4+X(I)*X(I)
                 S5=S5+Y(1,I)*Y(1,I)
210              CONTINUE

        T=N*S4-S1*S1
        S=(N*S3-S1*S2)/T
        B=(S4*S2-S1*S3)/T

        R=(N*S3-S1*S2)/
    2     (DSQRT(ABS(((N*S4-ABS(S1)**2))*(N*S5-ABS(S2)**2))))

        WRITE (5,260) S
260     FORMAT(//,1X,'SLOPE = ',D22.16)
        WRITE (5,280) B
280     FORMAT(//,1X,'INTERCEPT = ',D22.16)

        LITERL(2)='+'
        IF (ABS(B).NE.B) LITERL(2)='-'
        WRITE (5,290)
290     FORMAT(//,1X,'EQUATION FOR THE BEST LINEAR FIT IS : ')
        BBAR=ABS(B)
        WRITE (5,295) S,LITERL(2),BBAR
295     FORMAT(1X,'Y(X) =',D25.16,' * X ',A1,D25.16,///)
        WRITE (5,296) R
296     FORMAT(' LINEAR CORRELATION COEFFICIENT =',D22.16)
        STOP 'LINEAR... Execution completed'
        END
```