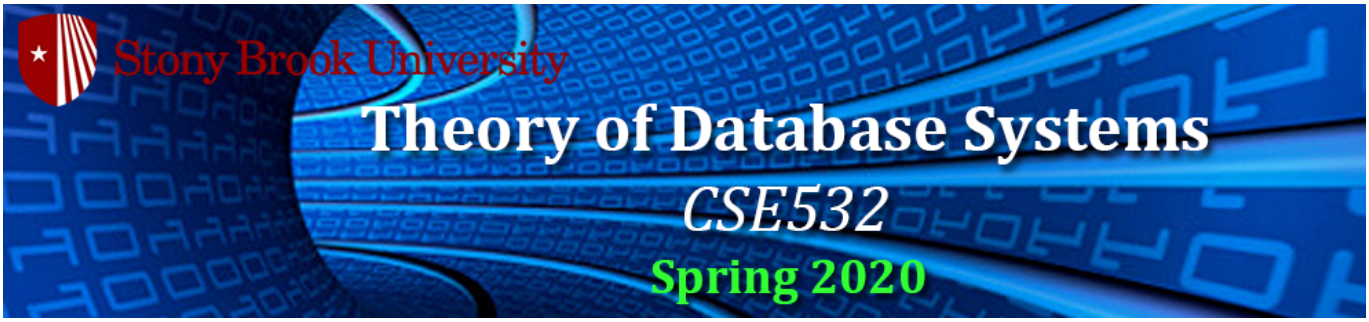

Home

CSE532

Home
Syllabus
Schedule
Homework
Textbooks
Examples
Instructor
Resources

Homework >

Homework2

Homework 2: Standard Deviation in JDBC and Stored Procedure

Update: one more parameter is added for the JDBC function to support automated grading. If you submit before or on Feb 23, you don't need to resubmit.

Due: March 2 midnight (11:59pm)
In this homework, you will write your own methods (in both JDBC and stored procedure) to compute standard deviation for employee's salary. You can validate your methods using SQL's stddev function.

Data: You will use the employee table in DB2 sample database. If you want to create the table by yourself, you need to create a table employee with this schema and load the sample data into the database first.

Standard deviation: In statistics, the standard deviation (**SD**, or **σ**) is a measure that is used to quantify the amount of variation or dispersion of a set of data values:

Let X be a random variable with mean value μ:

$$\mathrm{E}[X] = \mu.$$

Here the operator E denotes the average or expected value of X. Then the **standard deviation** of X is the quantity

$$\begin{aligned}\sigma &= \sqrt{\mathrm{E}[(X-\mu)^2]} \\ &= \sqrt{\mathrm{E}[X^2] + \mathrm{E}[-2\mu X] + \mathrm{E}[\mu^2]} = \sqrt{\mathrm{E}[X^2] - 2\mu\,\mathrm{E}[X] + \mu^2} \\ &= \sqrt{\mathrm{E}[X^2] - 2\mu^2 + \mu^2} = \sqrt{\mathrm{E}[X^2] - \mu^2} \\ &= \sqrt{\mathrm{E}[X^2] - (\mathrm{E}[X])^2}\end{aligned}$$

(derived using the properties of expected value).

In other words, the standard deviation σ (sigma) is the square root of the variance of X; i.e., it is the square root of the average value of $(X-\mu)^2$.

We will have two tasks for implementation the same function: one will be Java only (no stored procedure) using JDBC running on the client, and the other will be stored procedure only running on the server. Please DON'T write your JDBC code to execute your stored procedure.

Task 1: Write a JDBC program to compute the standard deviation of all salaries. Your program should run as "java -cp "path2jdbcdriver.jar" SalaryStdDev databasename tablename login password", where tablename is the tablename, and login and password are database login information. (We assume the database runs on localhost.)

Task 2: Write a SQL PL based stored procedure to compute the standard deviation of the salary. Your code should run as "db2 -td@ -f stddev.sql"

Performance considerations: Try to minimize the number of table scans. For example, one avg computation will scan the table once. Try to minimize system resources.

Submission: Please zip your codes (SalaryStdDev.java, stddev.sql), sample result (a file or screenshot), and a readme file on how to run your programs.

Please go to blackboard, and submit it under homework 2.

homework1sampledata.dat (5k)	Fusheng Wang, Feb 14, 2020, 9:09 AM	v.1
homework1tableschema.sql (0k)	Fusheng Wang, Feb 14, 2020, 9:09 AM	v.1