CONCORDIA UNIVERSITY
DEPARTMENT OF COMPUTER SCIENCE AND SOFTWARE ENGINEERING

SOEN 6611: SOFTWARE MEASUREMENT: SECTION AA
SUMMER 2017

PROJECT DESCRIPTION

Let x be a random variable that can take values from a finite data set x_1, x_2, ..., x_N, with each value having the same probability.

The **minimum**, m, is the smallest of the values in the given data set. (m need not be unique.)

The **maximum**, M, is the largest of the values in the given data set. (M need not be unique.)

The **mode**, o, is the value that appears most frequently in the given data set. (o need not be unique.)

The **median**, d, **is the middle number if N is odd, and is the arithmetic mean of the two middle numbers if N is even.**

The **arithmetic mean**, μ, is given by

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i.$$

The **standard deviation**, σ, is given by

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (x_i - \mu)^2}.$$

Let there be a system, called DESCRIPTIVE-STATISTICS, for finding m, M, o, d, μ, and σ. The system should take as input a random number of data values and output its descriptive statistics.

The purpose of the project is to create a **set of interrelated artifacts** for conducting certain measurements related to DESCRIPTIVE-STATISTICS. In the rest of the document, 'DESCRIPTIVE-STATISTICS' stands for the name of both the project and the product, unless otherwise stated. The work on DESCRIPTIVE-STATISTICS has been divided into a **collection of problems** to be solved.

DELIVERABLE 1

PROBLEM 1. [20 MARKS]

Using the **Goal-Question-Metric (GQM)** approach (or one of its extensions), present one goal specific to DESCRIPTIVE-STATISTICS and articulate 2N questions related to that goal, where N is the unit size. Discuss whether any metrics help answer those questions.

PROBLEM 2. [10 MARKS]

Using the given description, construct a **use case model** for DESCRIPTIVE-STATISTICS.

PROBLEM 3. [30 MARKS]

(a) Using the **use case points (UCP)** approach (or one of its extensions), provide an estimate of the effort for the project.

(b) Using Basic COCOMO 81, provide an estimate of the effort for the project.

(c) Comment on the difference in estimates using the UCP approach and COCOMO 81.

PROBLEM 4. [70 MARKS]

(a) Using **one** of the object-oriented programming languages (C++, C#, Java, or Python), implement DESCRIPTIVE-STATISTICS from **scratch**. This means, apart from the functions related to input and output and basic arithmetic, an implementation of DESCRIPTIVE-STATISTICS should **not** make use of any reuse mechanism (such as built-in functions, libraries, or APIs) provided natively by the programming language or otherwise. This may lead to **recursive implementation** of certain functions.

(b) Using input data that consists of (at least) 100 values, randomly distributed between 0 and 100, test DESCRIPTIVE-STATISTICS.

PROBLEM 5. [10 MARKS]

(a) Calculate the **cyclomatic number** (also known as cyclomatic complexity) of DESCRIPTIVE-STATISTICS.

(b) Comment on the qualitative conclusions that can be drawn with respect to the quantitative thresholds of the metric.

PROBLEM 6. [30 MARKS]

(a) Calculate the object-oriented metrics WMC, CF, and LCOM* for each of the classes of DESCRIPTIVE-STATISTICS. For WMC, assume that the weights are not normalized. Show your calculations in detail, both manually and using a tool, as applicable.

(b) Comment on the qualitative conclusions that can be drawn with respect to the quantitative thresholds of the respective metrics.

PROBLEM 7. [10 MARKS]

Calculate the Logical SLOC for DESCRIPTIVE-STATISTICS using the scheme followed by **Unified Code Count (UCC)**[1]. For specifics of source statement counts, download the UCC source and check.

NOTES

The source code for any purpose is part of the count. The test code, if any, should not be part of the count.

PROBLEM 8. [20 MARKS]

(a) Using **Scatter Plot**, carry out an analysis of the correlations between the data for Logical SLOC and WMC obtained from DESCRIPTIVE-STATISTICS.

(b) Using a **correlation coefficient**, carry out an analysis of the correlations between the data for Logical SLOC and WMC obtained from DESCRIPTIVE-STATISTICS.

NOTES

The submission must include **source code, test data, and documentation**.

DESCRIPTIVE-STATISTICS should aim for **generality**. The source code of DESCRIPTIVE-STATISTICS should aim for **readability** and **modularity**.

The documentation for DESCRIPTIVE-STATISTICS should be **organized**, and aim to be **clear and concise**.

The documentation for DESCRIPTIVE-STATISTICS should aim to be as **self-contained** as possible.

The documentation for DESCRIPTIVE-STATISTICS must **state explicitly any assumptions** made in the project.

[1] URL: http://csse.usc.edu/ucc_wp/ .

The documentation for DESCRIPTIVE-STATISTICS must include brief **descriptions of any algorithm(s)** used and the **rationale** for selecting the algorithm(s).

For clarity and to minimize the potential for ambiguity, the documentation for DESCRIPTIVE-STATISTICS must have a **glossary**.
The documentation for DESCRIPTIVE-STATISTICS must have **citations and references at appropriate places** corresponding to any non-original work in the project (that is, any work external to the unit). A comprehensive collection of resources on citing and referencing is available[2]. For example, ACM, APA, and IEEE provide standard formats for citing and referencing. It is important not to make claims that cannot be substantiated, and not to copy others' work verbatim regardless of whether it is cited. A copied work does not earn any credit.

[2] URL: http://library.concordia.ca/help/howto/citations.html .

DELIVERABLE 2

Given that thorough self-critique is the hallmark of any subject claiming to be a science, computing science cannot help but benefit from a retrospective analysis and evaluation.
— Niklaus Wirth

Present any critical decisions made during the project, and explain briefly why those decisions were critical.

Present any lessons learnt by doing the project.