SOEN 6611 (Software Measurement)

Concordia University
Department of Computer Science and Software Engineering

Project: METRICSTICS
Deliverable 1

Supervisor:
Prof. PANKAJ KAMTHAN

Students:

Team O[1]
Malay Sarjul Sheth (40225519)
Shreyans Sharma (40196896)
Snehil Sharma (40267285)
Joyal Rakeshkumar Shah (40226585)
Sanket Shinde (40233057)

October 1, 2023

Contents

List of Figures

List of Tables

List of Symbols and Abbreviations

Introduction

2.1 Objective

Descriptive statistics play a vital role in measurably condensing data, providing valuable understanding into data sets' central tendencies, variability, and frequency distributions. METRICSTICS, a blend word of "METRICS" and "STATISTICS", represents a system designed for the computation of key descriptive statistics, including minimum (m), maximum (M), arithmetic mean (μ), mean absolute deviation (MAD), and standard deviation (σ). This system is intended to intake a variable number of data values and produce their respective descriptive statistics. The data could be real-world, gained from comprehensive sources, or simulated with the help of random data producers.

The current extent of the project is limited to:

1. An input of arbitrary values from confined data set $(N_1, N_2, N_3, \ldots, N_M)$.
2. An output of any one of the following:
    - Descriptive Statistics: METRICSTICS produces statistics like minimum (m), maximum(M), mean (μ), mean absolute deviation (MAD), and standard deviation (σ).
    - Error Handling: The system must clearly define the error messages when it is incompetent in processing the data.

Problem 1

3.1 Problem Statement

Team needs to specify a goal using the Goal-Question-Metric (GQM) approach and formulate couple of questions per team member to reach at the defined goal within the given time span and figure out if there are any metrics that help us to answer those questions so that the required goal can be achieved.

3.2 Goal Formulation

The SMART goal for METRICSTICS is to develop an **efficient** and **accurate** system for calculating descriptive statistics from the given finite data sets within the upcoming four-month period.

3.3 SMART

SMART [2] principle consists of:

- **Specific**: The goal noticeably states what needs to be accomplished, which is the development METRICSTICS for descriptive statistics.
- **Measurable**: The objective can be measured by evaluating the efficiency and accuracy of the system.
- **Achievable**: Considering the project description, it is attainable to develop such a system.
- **Relevant**: The goal is directly related to project's purpose.
- **Time-bound**: A time frame of four months is specified for achieving the goal.

3.4 Questions Articulation & Metric Discussion

We articulate the following questions to guide the project:

1. How can we ensure the accuracy of descriptive statistics calculations in METRICSTICS?

   Metric: **Mean Absolute Error (MAE)** [3] - Calculate the MAE between the METRICSTICS' descriptive statistics that has been computed and the actual value (known) of the sample data sets provided as input. The lower the value of MAE, better is the correctness of the system.

2. What characteristics or capabilities should be prioritized in METRICSTICS development?

   Metric: **User Satisfaction Score** [4] - Gather feedback from users on the importance and satisfaction level of various features and functionalities present in the system by conducting surveys and analyzing the reviews section of the system. Prioritize those functionalities that have higher user satisfaction scores as they are most likely to be reused by users.

3. How can METRICSTICS enhance the performance of data sets with huge amount of data?

   Metric: **Processing Time** - Measure the time taken by METRICSTICS to calculate descriptive statistics for many data sets of various sizes. More optimizations, for instance using optimized algorithms or formulas, leads to less processing time.

4. What user requirements or expectations need consideration in METRICSTICS development?

   Metric: **User Requirement Compliance** - Analyze the extent to which METRICSTICS meets upon the agreed user requirements. A compliance percentage examines the degree to which system aligns with user expectations from the system.

5. What strategies can enhance the user interface and experience?

   Metric: **User Interface Usability Score** [5] - Conduct usability testing with users to evaluate how easy and intuitive it is for users to navigate or access the system using the given user interface. A higher usability score indicates a better user experience.

6. How will METRICSTICS handle missing or incomplete data?

   Metric: **Handling Efficiency** - Evaluate the system's effectiveness in handling missing or incomplete data by analyzing difference in processing time(key metric) between complete and missing data and resource utilization(key metric) like CPU usage, RAM(memory) usage and disk usage.

7. What are potential challenges or risks in METRICSTICS development?

   Metric: **Risk Severity Index** [6] - Assign a severity index to the risks that have been recognized based on the impact they can make on project goals or deliverable like exceeding project budget or crossing the project deadline. Put considerable efforts to reduce the impact of high-severity risks.

8. How can METRICSTICS ensure compliance with data privacy and security regulations?

   Metric: **Compliance Percentage** [7] - Conduct regular compliance audits to make sure that the METRICSTICS adheres to data privacy and security regulations. Compliance percentage indicates how much system is compliant with standards, rules, and laws.

9. What is the estimated development timeline, and how can it be maintained?

   Metric: **Task Completion Progress** [8]- Track the percentage of completed development tasks against the project schedule. Monitor adherence to upcoming tasks and milestones to make sure that the timeline is maintained and there are less high-severity risks.

10. How will the team coordinate effectively?

    Metric: **Task Collaboration Efficiency** - Assess the efficiency of team coordination by measuring the time taken to complete collaborative tasks and the frequency of successful task completions.

These metrics provide quantifiable measures for evaluating progress and performance related to each of the articulated questions, ensuring that project aims are met effectively.

Problem 2

4.1 Problem statement

Team needs to form a use-case model [9] for METRICSTICS where use-case scenarios are specified along with their description having actors involved in those use-cases.

4.2 Use Case Model

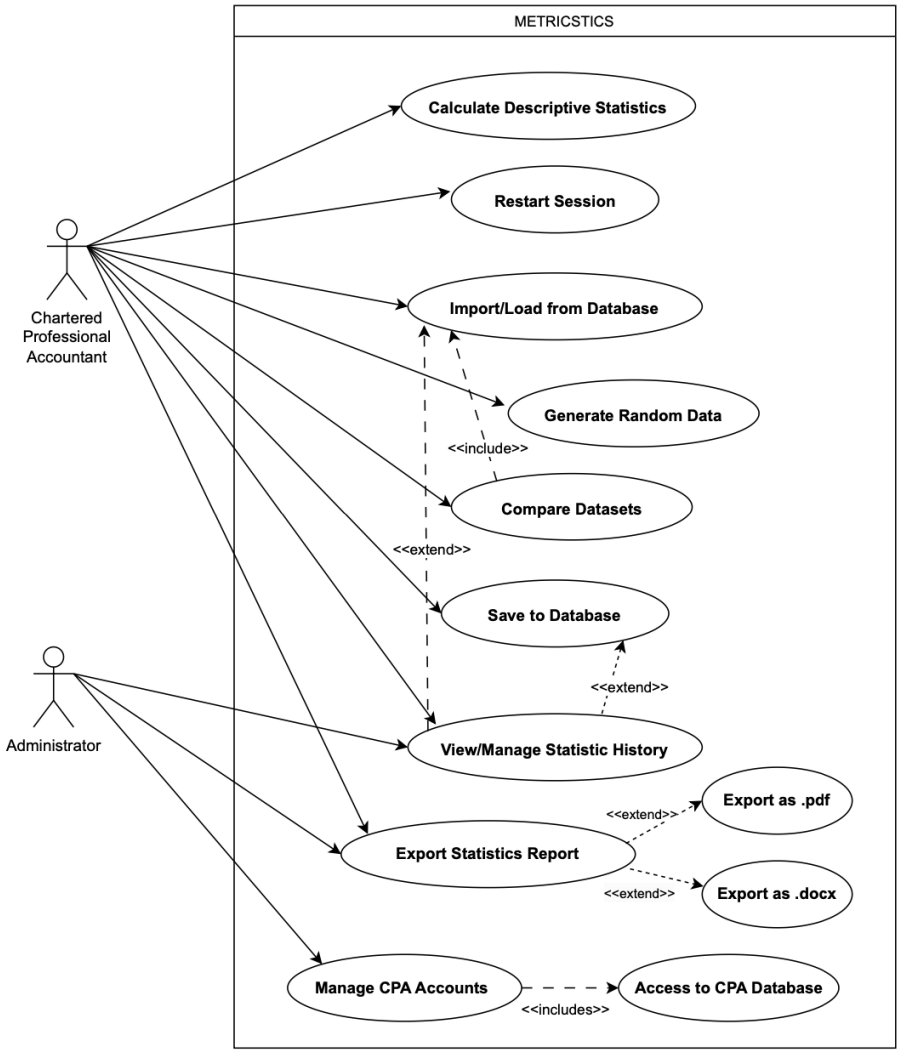

Figure 4.1: Use Case Model

4.3 Description Tables

Use Case (UC) ID	UC-1
Use Case Name	Calculate Descriptive Statistics
Primary Actors	CPA
Priority	High
Description	The CPA enters a collection of data values into METRICSTICS, and the system computes and displays the data's values for minimum, maximum, mode, median, mean, mean absolute deviation, and standard deviation.
Pre-conditions	The CPA is logged into the METRICSTICS system.
Post-conditions	The descriptive statistics for the entered data are displayed to the CPA.
Normal Flow	1. Chartered Professional Accountant (CPA) logs into METRICSTICS. 2. CPA selects the "*CalculateDescriptiveStatistics*" option. 3. CPA enters the data values. 4. The system computes and displays the descriptive statistics.

Table 4.1: Use Case-1, Calculate Descriptive Statistics

Use Case ID	UC-2
Use Case Name	Save to Database
Primary Actors	CPA
Priority	Medium
Description	The CPA saves a collection of data values along with the descriptive statistics that go along with it to a file for later use or study.
Pre-conditions	The CPA has calculated descriptive statistics for a set of data
Post-conditions	Data and associated statistics are saved to the database.
Normal Flow	1. CPA calculates descriptive statistics. 2. CPA selects the "*SaveToDatabase*" option. 3. CPA provides a name and description for the saved dataset. 4. System saves the data and statistics to the database.

Table 4.2: Use Case-2, Save to Database

Use Case ID	UC-3
Use Case Name	Import/Load From Database
Primary Actors	CPA
Priority	High
Description	The CPA loads into METRICSTICS for further analysis a previously saved set of data values and the related descriptive statistics.
Pre-conditions	The CPA is logged into the METRICSTICS system.
Post-conditions	The selected dataset and its associated statistics are loaded for analysis.
Normal Flow	1. CPA logs into METRICSTICS. 2. CPA selects the "$Import/LoadFromDatabase$" option. 3. CPA selects a previously saved dataset. 4. The system loads the data and associated statistics.

Table 4.3: Use Case-3, Import/Load From Database

Use Case ID	UC-4
Use Case Name	Restart Session
Primary Actors	CPA
Priority	Low
Description	The CPA can restart the whole calculation session from the start without saving any unnecessary data from the current session.
Pre-conditions	The CPA is in an active calculation session.
Post-conditions	The current session is reset, and no data is saved.
Normal Flow	1. CPA selects the "$RestartSession$" option. 2. The system clears the current session and resets it.

Table 4.4: Use Case-4, Restart Session

Use Case ID	UC-5
Use Case Name	Generate Random Data
Primary Actors	CPA
Priority	Medium
Description	The CPA specifies the range and number of data values, and METRICSTICS creates a set of random data values that may be tested or used in demonstrations.
Pre-conditions	The CPA is logged into METRICSTICS.
Post-conditions	A set of random data values is generated based on CPA's specifications.
Normal Flow	1. CPA logs into METRICSTICS. 2. CPA selects the "$GenerateRandomData$" option. 3. CPA specifies the range and number of data values. 4. The system generates random data based on the CPA's specifications.

Table 4.5: Use Case-5, Generate Random Data

Use Case ID	UC-6
Use Case Name	Compare Datasets
Primary Actors	CPA
Priority	Medium
Description	The system compares the descriptive statistics of two or more sets of data values to find similarities and differences.
Pre-conditions	The CPA has loaded multiple datasets for comparison.
Post-conditions	Comparison results are displayed to the CPA.
Normal Flow	1. CPA loads multiple datasets for comparison. 2. CPA selects the "$CompareDatasets$" option. 3. System analyzes and displays the comparisons.

Table 4.6: Use Case-6, Compare Datasets

Use Case ID	UC-7
Use Case Name	View/Manage Statistic History
Primary Actors	CPA, Administrator
Priority	Medium
Description	The CPA and the administrator can view/manage the descriptive statistics on data sets.
Pre-conditions	The CPA or administrator is logged into METRICSTICS.
Post-conditions	Descriptive statistics history is viewed or managed as required.
Normal Flow	1. CPA or administrator logs into METRICSTICS. 2. CPA or administrator selects the "$View/ManageStatisticHistory$" option. 3. The system displays the history or allows management actions.

Table 4.7: Use Case-7, View/Manage Statistic History

Use Case ID	UC-8
Use Case Name	Export Statistics Report
Primary Actors	CPA, Administrator
Priority	Medium
Description	The CPA and the administrator can export a given statistical report to an external data destination connected to the METRICSTICS system.
Pre-conditions	A statistical report is available for export.
Post-conditions	The statistical report is exported to an external destination.
Normal Flow	1. CPA or administrator selects the "$ExportStatisticsReport$" option. 2. CPA or administrator chooses the format for export (e.g., PDF, DOCX). 3. The system exports the report to the selected format.

Table 4.8: Use Case-8, Export Statistics Report

Use Case ID	UC-9
Use Case Name	Manage CPA Accounts
Primary Actors	Administrator
Priority	High
Description	The Administrator has access to manage the CPA accounts set up on the system.
Pre-conditions	The Administrator is logged into METRICSTICS.
Post-conditions	CPA accounts are managed as required.
Normal Flow	1. Administrator logs into METRICSTICS. 2. The administrator selects the "$ManageCPAAccounts$"$option$. 3. The system provides options for managing CPA accounts.

Table 4.9: Use Case-9, Manage CPA Accounts

Use Case ID	UC-10
Use Case Name	Access To CPA Database
Primary Actors	Administrator
Priority	High
Description	This use case simply gives the Administrator access to the CPA database connected with the METRICSTICS system to store details of its CPAs.
Pre-conditions	The Administrator is logged into METRICSTICS.
Post-conditions	The Administrator has access to the CPA database.
Normal Flow	1. Administrator logs into METRICSTICS. 2. The administrator selects the "$AccessToCPADatabase$" option. 3. The system grants access to the CPA database for management purposes.

Table 4.10: Use Case-10, Access To CPA Database

References

[1] Github Repository. https://github.com/MalaySheth/SOEN-6611-METRICSTICS.

[2] Pankaj Kamthan. The quality of goals in the gqm approach. https://users.encs.concordia.ca/~kamthan/courses/soen-6611/.

[3] Wikipedia. Mean absolute error. https://en.wikipedia.org/wiki/Mean_absolute_error.

[4] Retently. Customer satisfaction score (csat). https://www.retently.com/blog/customer-satisfaction-metrics/.

[5] Justin Mifsud. Usability metrics for satisfaction. https://usabilitygeek.com/usability-metrics-a-guide-to-quantify-system-usability/.

[6] Eclipse Suite. Risk severity. https://www.eclipsesuite.com/risk-severity/.

[7] Kasiddy Kelly. Compliance audit. https://www.techtarget.com/searchcio/definition/compliance-audit.

[8] Maxwell Matson. Task completion rate. https://www.playerzero.ai/advanced/kpi-guides/how-to-calculate-task-completion-rate.

[9] Visual Paradigm. Use case diagram at a glance. https://www.visual-paradigm.com/guide/uml-unified-modeling-language/what-is-use-case-diagram/.