**CS6.302 - Software System Development**

**Assignment 1 – SQL and NoSQL**

Due: 5 September 2025, 05:00 PM

Total Marks: 50

**NOTE**: This assignment is an individual submission, not a group activity. Evaluation will be conducted based on a fixed grading rubric (syntax, logic, input and output) and the marks are divided as per prescribed weightage in respective question. Inputs/output should fit the criteria mentioned in respective questions. Unless it is specified, all input/output criteria are open to interpretation. All questions in the assignment are self-explanatory. Do not reach us for any clarifications. If you are answering a question based on a certain assumption, please feel free to mention it as part of your README file.

**Submission Instructions:**

Please submit your code in Moodle. You are required to submit the assignment as <roll_number>A1.zip. Fort example, if you roll number 20162153, then your submission file should be 20162153_A1.zip. The submission ZIP file should contain two folders, Q1 & Q2 with each folder containing scripts and any supporting files associated with these questions. Please do not forget to include a README.TXT file to mention your assumptions, execution instructions or anything else in the ZIP. If you are using any LLM for this task, please declare your usage with all required details here - https://forms.office.com/r/754tAUacRk If you are found not mentioning about your LLM usage despite using one, you will be awarded '0'. You will be awarded '0' if your submission is found to be plagiarized with other submissions.

**Q1:** Load the supplied dataset[1] into a relational database and answer the following using SQL. **(30 Marks)**

a. Admission Funnel - An admission funnel is a stage-by-stage representation of how applicants progress through the selection process, showing the number who start, advance, or drop out at each stage until final admission. Write a single SQL query to present admission funnel of students in each stage of Admission and provide the average turn-around time between each stage for all students.

b. Pass and Fail Rate – Write a single SQL query provide pass rate by gender (Female/Male) in each stage, pass rate by age band (18-20, 21-23, 24-25) in each stage, pass rate by City in each stage.

c. Write a Stored Procedure that reads StudentID as input and summarizes the student performance by each stage of the test with other students from the same stage with dimensions gender, city, age. For example, if StudentID 20162153 is provided as input, the output should return his performance in each stage and also illustrate what is the mean pass/fail rate of other students by gender, city, age. You are open choose any other measure apart from mean.

[1]Dataset can be accessed from here. Details about the dataset are available here.

**Q2:** Load the supplied dataset[2] into a NoSQL database and answer the following using NoSQL **(20 Marks)**

a. Write NoSQL queries to identify daily average temperature per city by month, monthly average temperature per city, Identify hottest and Coldest cities from Overall Jan-Jun 2025.

b. Write NoSQL queries to identify top 5 hottest and coldest days nation wide, Compare City's temperature and weather data to determine if it is raining or not on a particular day, 7-day moving average (trend) for one given city.

[2]Dataset can be accessed from here. Details about the dataset are available here.

**_(-_-)_/ Happy -_- Programming!**