TIS3351 Advanced Database

TT1V

Assignment 1

Cheng Jia Pao 1191101533

Lecturer: Goh Chien Lee

Faculty of Computing and Informatics

Multimedia University

Trimester 2, 2021/2022

Name: Cheng Jia Pao
ID: 1191101533
Email: 1191101533@student.mmu.edu.my

Part 1

Entity-Relationship Diagram (ERD)

This section discusses the logical ERD of the operational data of the Faculty of Computing and Informatics at Multimedia University. The diagram is normalized to third normal form (3NF). There are seven tables in the database.
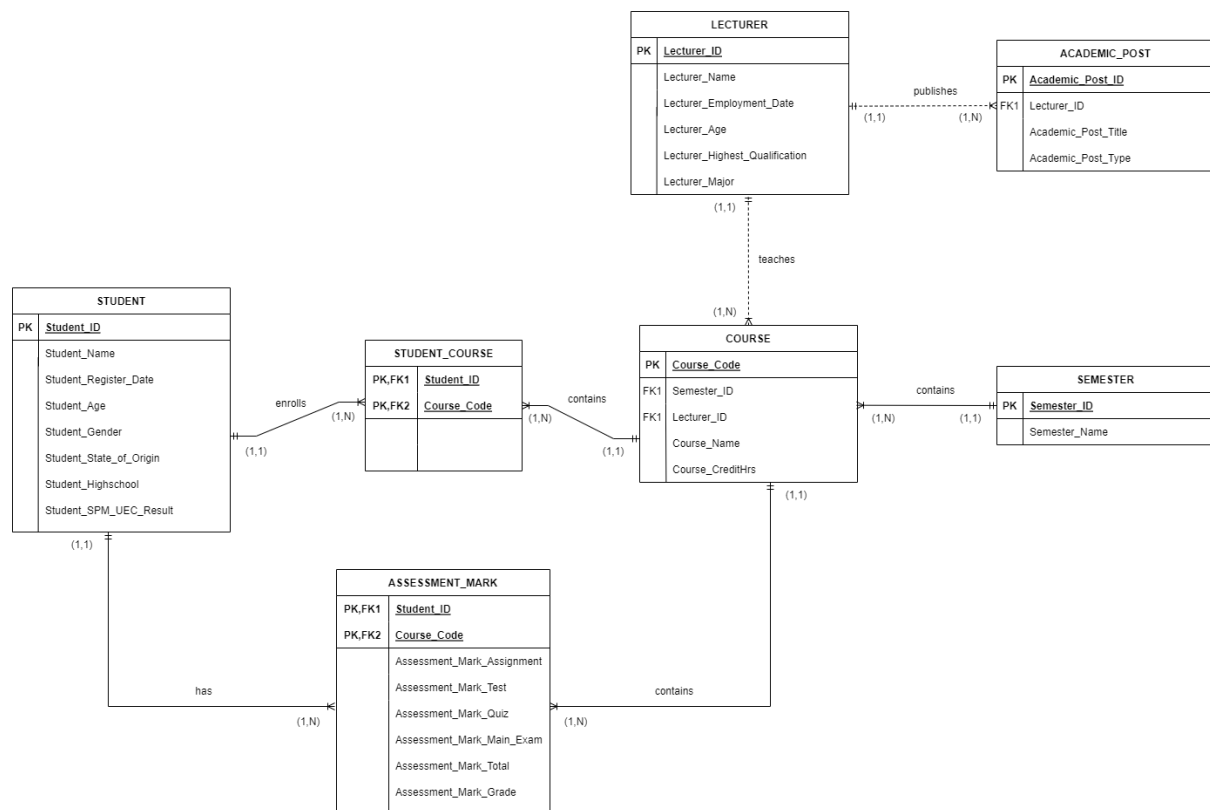

Figure 1.1: Logical ERD

Relational schema

There are seven entities in the designed database. Student_Course entity is a bridge entity to implement the M:N relationships between Student and Course. Course entity is also an associative entity to connect Lecturer table and Semester table.

Student (**Student_ID**, Student_Name, Student_Register_Date, Student_Age, Student_Gender, Student_State_of_Origin, Student_Highschool, Student_SPM_UEC_Result)

Lecturer (**Lecturer_ID**, Lecturer_Name, Lecturer_Employment_Date, Lecturer_Age, Lecturer_Highest_Qualification, Lecturer_Major)

Semester (**Semester_ID**, Semester_Name)

Course (**Course_Code**, Course_Name, Course_CreditHrs, *Semester_ID*, *Lecturer_ID*)

Assessment_Mark(***Student_ID***, ***Course_Code***, Assessment_Mark_Assignment, Assessment_Mark_Test, Assessment_Mark_Quiz, Assessment_Mark_Main_Exam, Assessment_Mark_Total, Assessment_Mark_Grade)

Academic_Post(**Academic_Post_ID**, Academic_Post_Title, Academic_Post_Type, *Lecturer_ID*)

Student_Course(***Student_ID***, ***Course_Code***)

Name: Cheng Jia Pao
ID: 1191101533
Email: 1191101533@student.mmu.edu.my

The relational schema is implemented in the SQL script *operational_db_creation.sql*. Sufficient data are inserted for each entity.

Design Assumption

This section discusses the design assumption made for the database model. The database is assumed made for the year 1 undergraduate student from the Faculty of Computing and Informatics at Multimedia University.

The Unified Examination Certificate (UEC) exam grades the results into A1, A2, B3, B4, B5, B6, C7, C8, and C9. Grade A indicates a score of distinction; Grade B is credit and Grade C is a pass. The students' UEC results will be recorded in terms of the number of grades A, B, and C as a total. Meanwhile for Sijil Pelajaran Malaysia (SPM) exam grade the results into A+, A, A-, B+, B, C+, C, D, and E. The students' SPM results will be recorded in terms of the number of grade A, B, C, D, and E as a total. The subject results of the students' that were failed or were absent from the subject would not be recorded.

It is also assumed that one academic year will have three semesters in which one semester only has three courses. Each course will only be available once in the academic year. While for assessment mark, it is assumed that there are only one assignment, one test, and one quiz.

Other than that, the business rule is listed down below to give a better insight into the ERD.

Business Rule

1. Each Semester contains many Courses.
   Each Course belongs to one and only one Semester.
2. Each Student enrolls in many Courses.
   Each Course is enrolled by many Students.
3. Each Course contains many Assessment Marks.
   Each Assessment Mark belongs to one and only one Course.
4. Each Student has many Assessment Marks.
   Each Assessment Mark belongs to one and only one Student.
5. Each Lecturer publishes many Academic Posts.
   Each Academic Post is published by one and only one Lecturer.
6. Each Lecturer teaches many Courses.
   Each Course is taught by one and only one Lecturer.

In this database model, there will be thirty Students, six Lecturer, and nine Courses where there are three Courses for one semester. Assumed that one Course is only available in one and only one of the three semesters within the academic year. In each Semester, each Student is mandated to take all the available Courses. Logically, one Assessment Mark is generated for one Course; since there are nine Courses in one academic year, each Student will have nine Assessment Marks which made up to 270 records of Assessment Marks in total. One Lecturer will publish one to many Academic Posts since it is assumed that one Academic Post will be published for each degree level that the Lecturer has completed. One lecturer will teach one to more courses but each Course will only be taught by one Lecturer.

Name: Cheng Jia Pao
ID: 1191101533
Email: 1191101533@student.mmu.edu.my

Part 2

Snowflake Scheme ERD

This section discusses the ERD of the data warehouse. The ERD is a snowflake scheme with one fact table and four dimensions. One of the dimensions, Lecturer has a dimension named Academic Post.
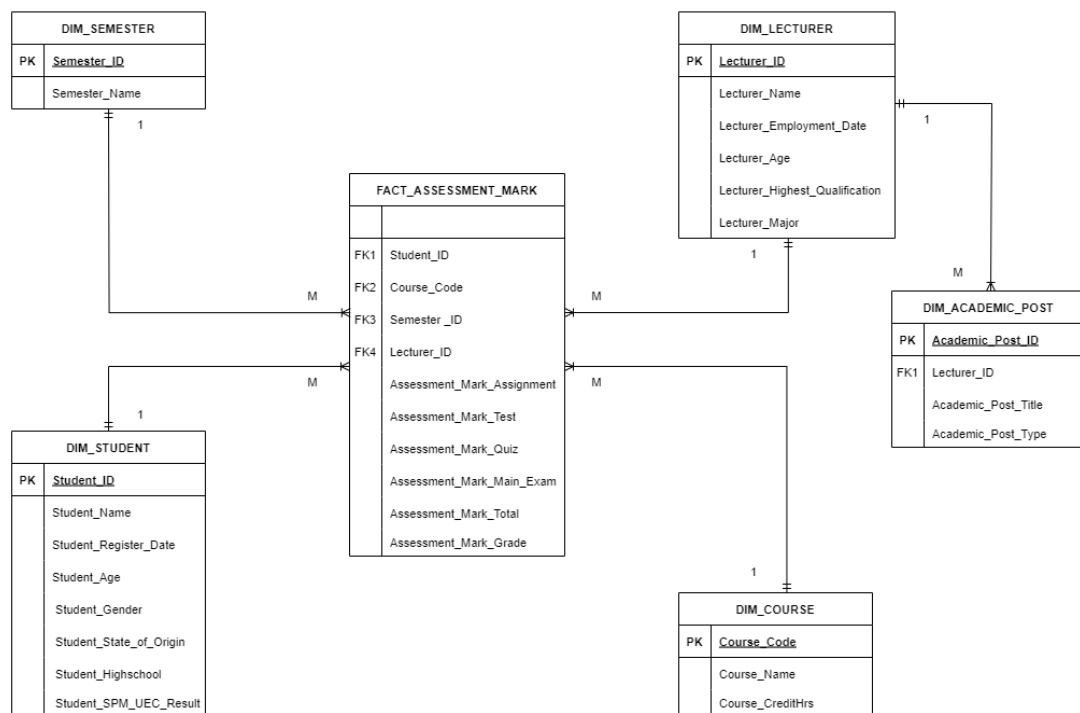

Figure 2.1: Snowflake Schema ERD

Relational schema

There are six entities in the designed data warehouse. Assessment Mark is the fact; while Student, Semester, Course, and Lecturer are the dimensions of Assessment Mark. Lecturer dimension has a dimension called Academic Post.

Dim_Student (**Student_ID**, Student_Name, Student_Register_Date, Student_Age, Student_Gender, Student_State_of_Origin, Student_Highschool, Student_SPM_UEC_Result)

Dim_Semester (**Semester_ID**, Semester_Name)

Dim_Course (**Course_Code**, Course_Name, Course_CreditHrs)

Dim_Lecturer (**Lecturer_ID**, Lecturer_Name, Lecturer_Employment_Date, Lecturer_Age, Lecturer_Highest_Qualification, Lecturer_Major)

Dim_Academic_Post(**Academic_Post_ID**, Academic_Post_Title, Academic_Post_Type, *Lecturer_ID*)

Name: Cheng Jia Pao
ID: 1191101533
Email: 1191101533@student.mmu.edu.my
Fact_Assessment_Mark(***Student_ID***, ***Semester_ID***, ***Course_Code***, ***Lecturer_ID***, Assessment_Mark_Assignment, Assessment_Mark_Test, Assessment_Mark_Quiz, Assessment_Mark_Main_Exam, Assessment_Mark_Total, Assessment_Mark_Grade)

The data warehouse is implemented in the SQL script *dw_creation.sql*. Data are extracted from the operational database and loaded into the data warehouse by semester. Data for semester 1 can be found in *tri1.sql*; Data for semester 2 can be found in *tri2.sql*; Data for semester 3 can be found in *tri3.sql*.

Design Assumption

In the creation of the data warehouse for the year 1 undergraduate student from the Faculty of Computing and Informatics at Multimedia University, it is assumed that the data warehouse is used to store historical data from the operational data. Assumed that one fact table is sufficient for the data warehouse to record the assessment mark of the students.

Name: Cheng Jia Pao
ID: 1191101533
Email: 1191101533@student.mmu.edu.my

Part 3

This section includes five queries implemented in the data warehouse to help the President of MMU to gain insight into the performance of the students and lecturers of MMU. Each query generates a csv file which is then used in Excel to generate a graphical chart to give a better visualization of the result.

Query 1

This query is used to calculate the number of grades achieved by the students by subject. A bar chart gives a better distribution of the result.
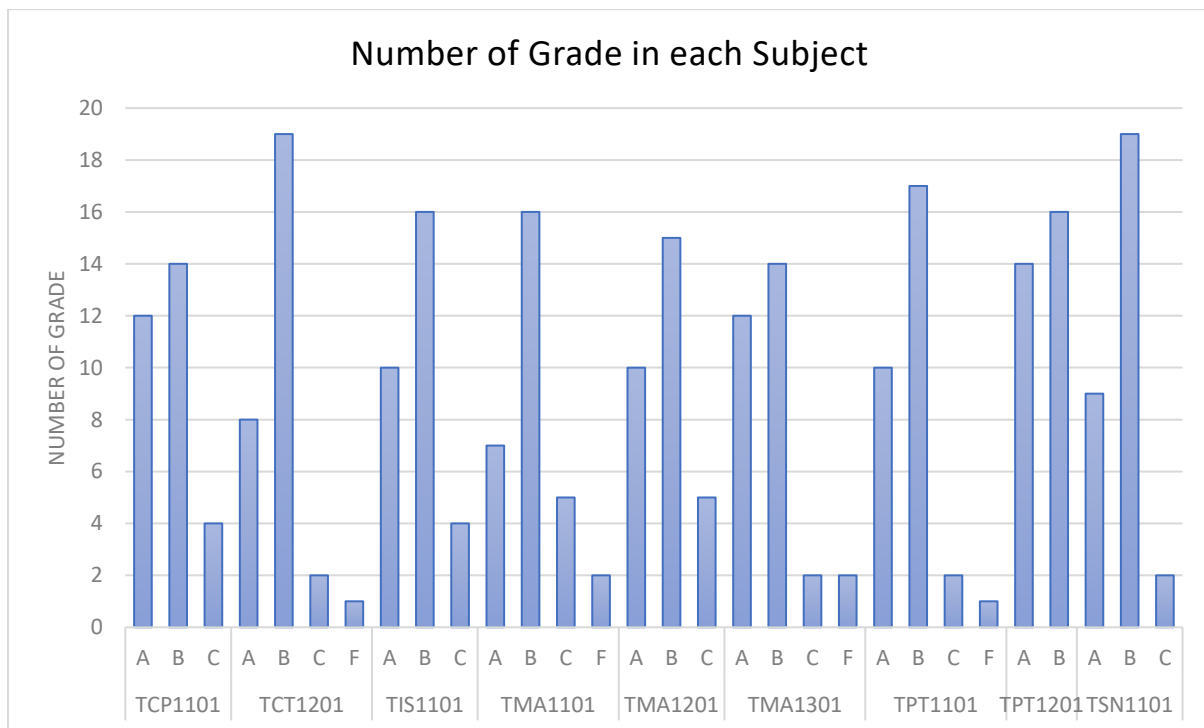

Figure 3.1: Number of grades of students by subject

Name: Cheng Jia Pao
ID: 1191101533
Email: 1191101533@student.mmu.edu.my

Query 2

This query helps the President of MMU study the lecturers' teaching performance. It shows the number of students who fail the subject taught by the respective lecturer. This bar chart only shows the name of the lecturer and the respective subject which exists a student fails. The lecturers of MMU are performing well as the number of students who fail are very less.
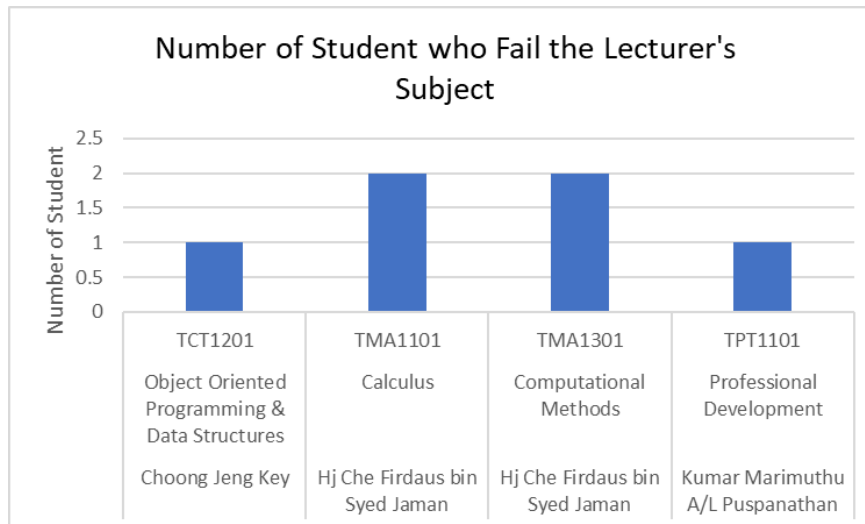

Figure 3.2: Number of students who fail Lecturer's subject

Query 3

This query helps the President of MMU to know the students' performance. It calculates the average mark of each subject. The graph shows that the students have a high average score indicates that the students are doing well in the subject.
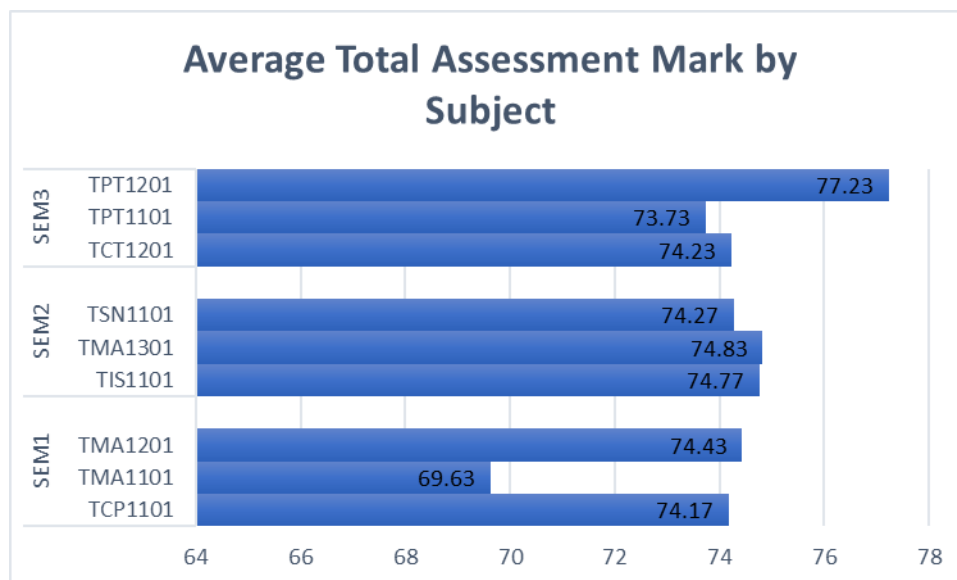

Figure 3.3: Average total assessment mark of each subject

Name: Cheng Jia Pao
ID: 1191101533
Email: 1191101533@student.mmu.edu.my

Query 4

This query displays the mode and the min of the total assessment mark of each subject. This helps to study the difference in the total assessment mark achieved by the students.
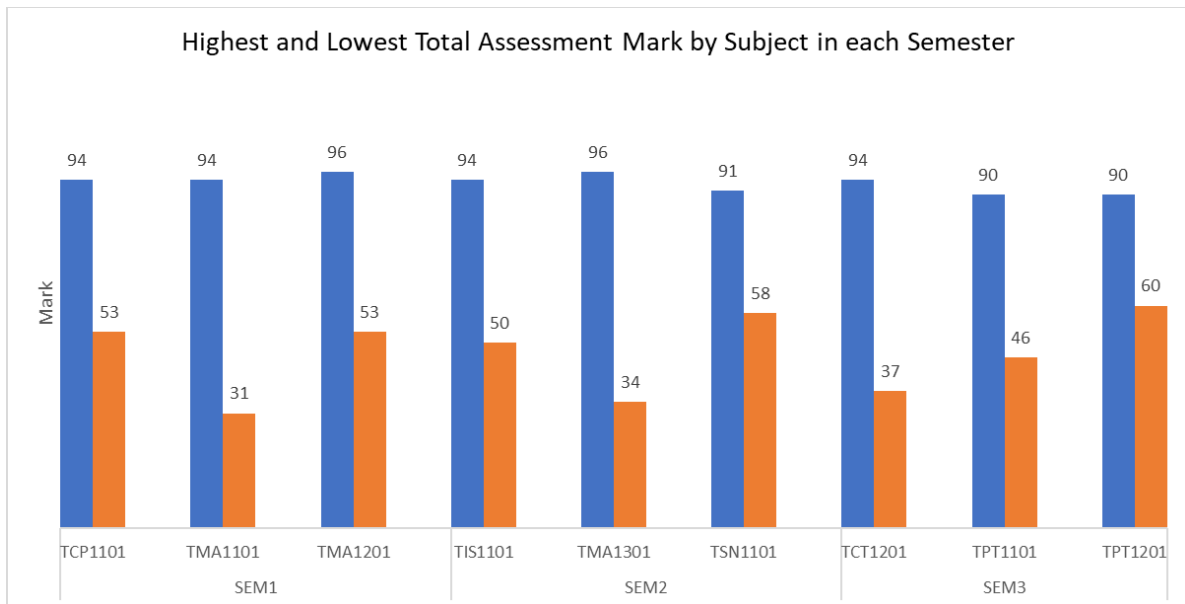

Figure 3.4: Highest and lowest of the total assessment mark of each subject

Query 5

This query helps the President of MMU to find the lecturer who has the highest number of students who got an A in the subject taught by the lecturer. The president can use this data to perform further actions such as increasing the lecturer's bonus.
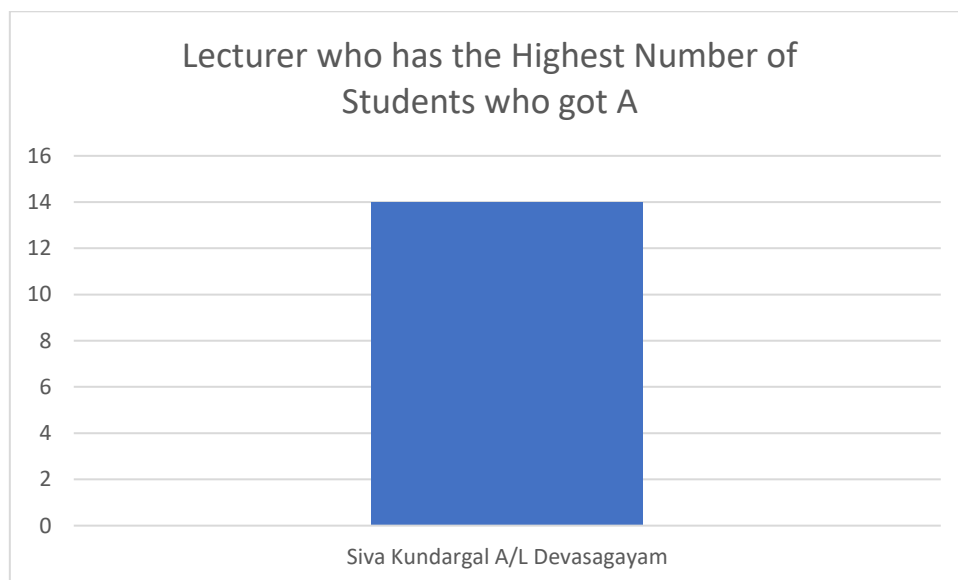

Figure 3.5: Lecturer who has the highest number of students who got A