Softwarica College of IT & E-Commerce

ST4005CEM Database System

in collaboration with

Assignment Brief 2021

Module Title: **Database Systems**	Ind/Group **Ind**	Cohort **May 2021**	Module Code: **ST4005CEM**
Coursework Title (e.g. CWK1) Assignment			Hand out date: TBD
Module Leader Bikash Ghimire			Due date: TBD
Estimated Time (hrs): Word Limit*: 750	Coursework type: Individual / Practical/Viva		% of Module Mark 100%
Submission arrangement online via Softwarica Moodle: Upload through assignment links Mark and Feedback date: Within 3 weeks of final assignment submission Mark and Feedback method: Rubric marks and comments			

Module Learning Outcomes Assessed:

1. Understand the sources of data in society, how to collect such data, the problems that might occur in collecting and storing such data and the basic statistics that are routinely performed on it.
2. Model databases using techniques such as normalization, entity relationship diagrams and a document-based approach
3. Create, populate and perform basic access approaches on a database using SQL Database Management System (DBMS).
4. Analyze different types of data using tools such as R or Python

Task and Mark distribution:

1. Normalization (10%)
2. ER Diagram (15%)
3. SQL (25%)
4. Data Visualization (25%)
5. Viva (25%)

Notes:

1. You are expected to use the CUHarvard referencing format. For support and advice on how this students can contact Centre for Academic Writing (CAW).
2. Please notify your registry course support team and module leader for disability support.
3. Any student requiring an extension or deferral should follow the university process as outlined here.

4. The University cannot take responsibility for any coursework lost or corrupted on disks, laptops or personal computer. Students should therefore regularly back-up any work and are advised to save it on the University system.
5. If there are technical or performance issues that prevent students submitting coursework through the online coursework submission system on the day of a coursework deadline, an appropriate extension to the coursework submission deadline will be agreed. This extension will normally be 24 hours or the next working day if the deadline falls on a Friday or over the weekend period. This will be communicated via email and as a Softwarica Moodle announcement.

Assignment Brief

Himalaya Tours

A Himalaya tours wants to get its database developed. Requirements are following.

The system needs to keep track of people. For each person, it records his/her address. Each address consists of country, province/state, city, street, street number, zip code/post code, and a list (possible empty) of phone numbers and a list of email address.

Each person in the database can be an old customer (have taken a tour of the company), a current customer (is booked to take a tour or is on a tour right now), a tour guide, an employee (works for the tour company), or any mixture of these (for instance an employee can take a tour and so can be a customer as well). The sex and age of each person must also be recorded.

The system also keeps track of all tours, past and future. Each tour has a unique number, itinerary, guide (at least one, but may be more than one), its status (completed, in-progress, in-the-future), and the list of participants. The amount paid by each person for the tour is also recorded. Each person is provided with an itinerary that consists of list of the dates the tour covers and for each date it includes the place of breakfast, the place of lunch, the place of dinner, and the accommodation and room. The type of room whether shared or private should also be mentioned on the itinerary. Each day in the itinerary also includes a simple English description of the activities during that day.

Note that the information provided here may not be sufficient to build complete database for the Himalaya Tour. Please make reasonable assumptions about the scenario. You have to make explicit in your assignment solution about various assumptions you have made in provided scenario.

Stock Data

Given sample data is for visualization purpose.

	Stock1	Stock2	Stock3	Stock4
Day1	185.74	1.47	1605	95.05
Day2	184.26	1.56	1580	97.49
Day3	162.21	1.39	1490	88.57
Day4	159.04	1.43	1520	85.55
Day5	164.87	1.42	1550	92.04
Day6	162.72	1.36	1525	91.7
Day7	157.89	NA	1495	89.88
Day8	159.49	1.43	1485	93.17
Day9	150.22	1.57	1470	90.12
Day10	151.02	1.54	1510	92.14

Task1: Normalization (10%)

Reflect understanding of dependencies and normalization.

Task 2: Entity Relationship diagram (15%)

Draw an entity relationship diagram for this system. Identify all relevant entities, attributes, suitable attribute types, primary keys, foreign keys, and cardinalities between entities. This should be normalized up to 3rd Normal Form. You should document all the steps you took to come up with this normalized ER diagram.

Task 3: SQL (25%)

Create a database using MySQL- with primary keys, foreign keys, and other attributes mentioned for each entity using proper constraints. Input some significant data in each table showing your understanding of the scenario. Write one SELECT and one JOIN query

Task 4: Data Visualization (25%)

Create given sample data of Stock in desired format and import to modeling tool. Use these data and perform the simple linear regression plot between stock1 and stock4 in R and also perform the correlation between them. Fit the linear regression and show the summary of the model

Task 5: Viva (25%)

How and what is done in all above task need to be summarized during Viva session

Marking Rubrics

T1: Normalization (10%)

0	1	2	3	4	5
Failed to submit the task within deadline. Normalization is not done	Only up to 1NF is shown or explained	Complete 1NF is done and all functional Dependency is Identified and also Single value attribute Is maintained	Up to 2NF is shown Or explained Partial dependency is Identified and removed	Up to 3NF is shown Or explained Transitive or non-key Dependency is Identified and removed	Excellent Normalization With all dependency Identified and removed Shown in tabular Format with clear Model and all relevant Relation and respective attributes

T2: Entity Relationship diagram (15%)

0	1	2	3	4	5
No ER diagram	ER diagram with only Few relevant entities, Attributes, their type, Primary key and Foreign key is identified	ER diagram with relevant Entities and their cardinalities are identified and depicted. The Process of normalization is not documented.	Good ER diagram With relevant Entities and their cardinalities are Identified and depicted. The Process of Normalization is not documented.	Good ER diagram With relevant Entities and their cardinalities are Identified and depicted. The Process of Normalization is documented.	Excellent ER diagram With relevant Entities and their cardinalities are Identified and depicted. The Process of Normalization is Also justifiably documented.

T3: SQL (25%)

0	1	2	3	4	5
No SQL Query	Only create table Query without Keys and relevant Attribute	Create table query With keys and Relevant attribute	Create table query With keys and Relevant attribute. Proper datatype And length of attribute along with Output	Create table query With keys and Relevant attribute. Proper datatype And length of attribute along with output. Insert query with Significant data	Create table query with keys and Attribute. proper Datatype and length Of attribute along with Output. Insert query With significant data. Shown one select And one join query

T4: Data Visualization (25%)

0	1	2	3	4	5
No any visualization	Only import of Sample data into Modeling tool	Performing linear Regression plot	Performing linear Regression plot Along with correlation	Performing linear Regression plot Along with correlation And fit the linear regression	Summary of the model

T5: Viva (25%)

0	1	2	3	4	5
Viva not given	Only few tasks explained In viva session	Only few tasks explained And little basic concept In viva session	Good summary of Completed task with how and what Done in assignment	Excellent summary of Completed task With how and what Done in assignment	Excellent summary of Completed task with How and what done In assignment with Proper output justification