**XJTLU Entrepreneur College (Taicang) Cover Sheet**

|  |  |  |
| --- | --- | --- |
| Module code and Title | CPT103TC: Introduction to Database | |
| School Title | School of AI and Advanced Computing | |
| Assignment Title | Assessment Task 001 (CW) | |
| Submission Deadline | **15th May 2022 at 5:00 PM** | |
| Final Word Count | NA | |
| If you agree to let the university use your work anonymously for teaching and learning purposes, please type **“yes”** here. | | **Yes** |

I certify that I have read and understood the University’s Policy for dealing with Plagiarism, Collusion and the Fabrication of Data (available on Learning Mall Online). With reference to this policy I certify that:

* My work does not contain any instances of plagiarism and/or collusion.  
  My work does not contain any fabricated data.

**By uploading my assignment onto Learning Mall Online, I formally declare that all of the above information is true to the best of my knowledge and belief.**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Scoring – For Tutor Use** | | | | | | | | | |
| **Student ID:** | | | |  | | | | | |
| **Group No:**  **(please provide your group number here to locate your group work report)** | | | |  | | | | | |
|  | | | | | | | | | |
| **Stage of Marking** | | **Marker**  **Code** | **Learning Outcomes Achieved （F/P/M/D）**  **(please modify as appropriate)** | | | | **Final**  **Score** | | |
| **A** | **B** | **C** | **D** |
| 1st Marker – red pen | |  |  |  |  |  | **Group** | **Individual** | **Total** |
|  |  |  |
| Moderation  – green pen | | **IM**  **Initials** | The original mark has been accepted by the moderator (please circle as appropriate): | | | | Y / N | | |
|  | Data entry and score calculation have been checked by another tutor (please circle): | | | | Y | | |
| 2nd Marker if needed – green pen | |  |  |  |  | |  | | |
| **For Academic Office Use** | | | **Possible Academic Infringement (please tick as appropriate)** | | | | | | |
| **Date**  **Received** | **Days late** | **Late Penalty** | **Category A** | | Total Academic Infringement Penalty (A,B, C, D, E, Please modify where necessary) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | | | | |
|  |  |  | **Category B** | |
| **Category C** | |
| **Category D** | |
| **Category E** | |

**Students**

**(Please modify where necessary)**

The assignment must be typed in an MS Word document and submitted as a pdf via Learning Mall Online to the correct dropbox. Only electronic submission is accepted and no hard copy submission.

All students must download their file and check that it is viewable after submission. Documents may become corrupted during the uploading process (e.g. due to slow internet connections). However, students themselves are responsible for submitting a functional and correct file for assessments.

**Assessment 001: Individual Work (50%)**

**Due:** May 15th, 2022 @ 17:00

**Weight: 60%** (30% individual work+30% group work)

**Maximum Marks:** 100

The course work will be assessed for the following learning outcomes:

1. Demonstrate a basic understanding of the design of databases.
2. Show a fundamental grounding in the operation and usage of database management systems including "hands-on" experience of a basic database management system.
3. Demonstrate in-depth knowledge of the database language, SQL.
4. Show understanding of the legal processes and implications of creating and maintaining information systems.

**Individual Assessment Tasks (50%)**

Each student should work on the same database they created in the group and provide following tasks in the individual report.

1. Data Visualization**(10%)**
2. SQL queries that can be run against the database **(15%)**
3. Relational algebra queriesthat can be run against the database **(10%)**
4. Presentation and question answering **(15%)**

You will be required to write a brief report for each component and note down your process, thoughts, and assumptions made.

**Marking Criteria**

Please see the assessment rubric at the end of this document outlining the criteria for individual work assessment. Please append the assessment rubric at the end of your individual report for grading purpose.

**Q1: Data Visualization (10%)**

Each team member writes a script to quickly explore data using visualization such as Pi, bar, line, scatter charts etc, For example:

* A bar chart for displaying products at the horizontal axis and quantity sold at the vertical axis to quickly explore which product is selling more.

You need to think carefully to craft a question statement in English for visualization that should precisely convey the information to be useful for the viewers. The visualization should answer one question, variable types used should be suited for the type and scale of data they represent, and legend should describe every variable type used in the visualization.

In your report provide the visualization question, script or procedure that has been used to create the visualization and a screenshot of the visualization that has been created upon executing the script.

**Q2: SQL Queries (15%)**

Write five different queries on the database you have created in the group work to showcase different SQL skills. Save your queries with English specifications in a text file with extension sql (for example SqlScript.sql) and submit it along with the report.

Challenging queries would amount to higher marks. The complexity of a query will be measured by the number of joins, select conditions and Group By clause, where a select condition counts as a ‘1’ and a join count as a ‘2’ and Group By condition count as ‘3’ (so, a SQL query that probably involved two select conditions, one join condition, and Group By clause will be counted as the complexity of ‘4’).

Below are the examples of queries statement in English for our hypothetical automobile company for your reference.

1. Show sales trends for various brands over the past 3 years, by year, month, and week. Then break these data out by gender of the buyer and then by income range.
2. Suppose that it is found that transmissions made by supplier Getrag between two given dates are defective. Find the VIN of each car containing such transmission and the customer to which it was sold. Suppose the defective transmissions all come from only one of Getrag’s plants if your design allows.
3. Find the top 2 brands by dollar amount sold in the past year.
4. Find the top 2 brands by unit sales in the past year.
5. Find those dealers who keep a vehicle in inventory for the longest average time.
6. In what month(s) do convertibles sell best?

**Q3: Relational Algebra Queries (10%)**

Write down five relational algebra queries that can be evaluated with your schema definitions. You can either write equivalent relational algebra queries of previous SQL questions or write different queries. In any case, you should state the English specification of the query, as well as state the query in relational algebra expression.

The complexity of a query will be measured by the number of joins, select conditions and aggregate function, where a select condition counts as a ‘1’ and a join count as a ‘2’ and aggregate function count as ‘3’ (so, a relational algebra expression that probably involved two select conditions, one join condition, and an aggregate function will be counted as the complexity of ‘4’).

**Project Presentation (15%)**

Demonstrate your end to end system, each team member should select a part of the database and interface to be demonstrated followed by a 3-4 minute question-answer session for each team member. The presentation should be no longer than 30 minutes for each project.

**Report Submission Guidelines**

Each student should submit individual work along with the script files at LMO. Everything should be packed in a zip file.

In the zip file:

1. A document called “[student ID].pdf”, which is your individual report. Please append the individual work rubric at the end of the report for grading purposes.
2. You should submit QueryScript.sql that can be executed against the database you have created in the group

If you are making any changes in the group work database in any one of the following files, please submit the complete updated scripts.

1. Create table statements saved as ddl.sql
2. Insert statements saved as data.sql

**Individual work rubric**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Category** | **4** | **3** | **2** | | **1** | **0** | Points | Weight | Marks Received =Points\*weight |
| **Data Visualization (10%)** | Creative visualization that conveys precise information to be useful for viewers, variable types used are suited for the type and scale of data they represent, and legend describes every variable type used in the visualization | Visualization conveys some confusing information, most of the variable types used are suited for the type and scale of data they represent, and legend describes every variable type used in the graphics | Visualization conveys too little or too much information, most of the variable types used are suited for the type and scale of data they represent, and legend describes every variable type used in the graphics | | Poorly defined visualization question, some variables do not make logical sense, either there is no legend or it does not describe any of graphic variable | Visualization was not created |  | **2.5** |  |
| **Querying (15%)** | Consistency between the English specification of the query, and the SQL implementation Demonstrated mastery in SQL skills by 4-5 complex queries | Consistency between the English specification of the query, and the SQL implementation  Adequately demonstrated SQL skills by creating 2-3 complex queries using joins and data filter options | Some inconsistencies between the English specification of the query, and the SQL implementation  Adequately  Created 1-2 complex queries by using different data filter options and joins | | Inconsistency between the English specification of the query, created only basic queries to retrieve data from a single table. Use of data filter and joins were not evident | Relevant queries were not created |  | **3.75** |  |
| **Relational Algebra (10%)** | Consistency between the English specification of the query, and the relational algebra implementation created 4-5 complex queries | Consistency between the English specification of the query, and the relational algebra implementation created 2-3 complex queries | Some inconsistency between the English specification of the query, and the relational algebra implementation created 1 complex query | | Inconsistency between the English specification of the query, created only basic, use of data filter and joins were not evident | Relevant queries were not created |  | **2.5** |  |
| **Presentaion(6+8=15%)** | The individual presented a portion of each of the following:   * Entities * Attributes * Relationships * tables | The individual presented 3 of the following:   * Entities * attributes * relationships * tables | The individual presented 2 of the following:   * Entities * attributes * relationships * tables | | The individual presented 1 of the following:   * Entities * attributes * relationships * tables | The individual did not participate |  | **1.5** |  |
| Individuals answered 90-100% of the questions using clear, concise rationale. | Individuals answered 80-89% of the questions using clear, concise rationale. | Individuals answered 70-79% of the questions using a clear, concise rationale | | Individuals answered 1-69% of the questions | The individual did not participate |  | **2** |  |
| **Total Marks = 50** | | | | **Marks Obtained** | | |  | | |