

**ISYS2038 Database Design and Development  
Semester 1, 2024  
Individual Assignment – Specifications**



**Released Date:** Tuesday 4 June, 10.00am  
**Due Date** : Wednesday 12 June, 11.59pm  
**Mark** : 50% of the total assessment of this course

## **1) Background**

In this assignment, you are required to demonstrate and apply the concepts and knowledge covered in Week 1 – Week 11. For the report, assume that you are employed as a business analyst at Best Innovative Solution (BIS) Pty Ltd. You are assigned to investigate the questions from a new client and provide the solution.

The dataset about Stella Pty. Ltd. has been collected, and you are given the following two files to apply data analytics:

- dataset (StellaCustomers.xlsx)
- header description (StellaCustomersDataset-header-description.txt)

These two files are available for you to download on *Canvas | Assignments | Assessment Task 3: Individual Final Assessment | Individual Assignment – Case Study, Specifications and Dataset*. Further information about the dataset is listed below:

- The dataset and header description are stored in two separate files.
- The dataset consists of binary, categorical and numerical data.
- There are seven different attributes and 1000 cases (or instances).
- There are some missing values in the dataset.
- The dataset covers mainly the aspect of customer-related data.

## **2) Questions from the new client**

***You need to:***

- (a) answer all the following questions based on the case study***  
***(b) include the word count on the cover page of your report***

### ***Question 1: Data Modelling***

- Provide an Entity Relationship Diagram (ERD). For each entity, provide at least one attribute/ field with a maximum of two attributes (excluding primary key).
- Provide a Relational Model and show the most appropriate attributes/ fields, primary keys, and foreign keys. In addition, include the data type for each attribute/ field listed in each table/ entity.
- You should apply the third normal form (3NF) and show consistency in the use of notation.
- If you discover while drawing the diagrams that the narrative of the case study is incomplete, then provide justifications, assumptions, or reasonable explanations to complete the case study along with your diagrams. You could include the justifications, assumptions, or reasonable explanations under the Appendix section.

### **Question 2: SQL Scripts**

- Based on data modelling from Question 1, create 3 queries that have the following requirements. Note that each of the query can cover one or more of the requirements.
  - One of the 3 queries should have a calculation.
  - One of the 3 queries should make use of Group By.
  - One of the 3 queries should include a scalar function (one that returns a value).
  - One of the 3 queries should demonstrate Nested Query (could be either standard or correlated sub-query).
  - One of the 3 queries should demonstrate joining tables.
  - One of the 3 queries should demonstrate the creation and call of Function.
- For each query,
  - Explain and justify its business purpose and business value or impact.
  - Provide the SQL script.
  - Explain the design of SQL script.
  - Make use of table(s) with sample data to show the potential result/outcome of the query.

### **Question 3: Data Analytics with Orange**

Discuss how you have dealt with the missing values and predicted the spending score of customers for Stella Pty. Ltd. Include the screenshots of using Orange to support your explanations. In addition, please explain how your results from visualisation(s) and prediction model(s) could help Stella Pty. Ltd. expand its business.

### **Question 4: Database Administration Approaches**

Describe two potential impacts on Stella Pty. Ltd. If data is not appropriately protected. For each type of impact, discuss a recommended security feature of a database management system that Stella Pty. Ltd. could consider for data protection.

### **Question 5: Big Data and Analytics**

Suggest two applications of big data and analytics that can assist Stella Pty. Ltd. in understanding its business. Discuss the benefits where each of the suggested application would deliver to the company.

## **3) Report**

**Word Limit:** 1500 – 2000 words +/- 10% (excluding Table of Content, Appendix, figures, tables, screenshots and references, anything beyond word limit will not be marked)

**Font Size:** 11pt or 12pt

**Font Style:** Calibri or Times New Roman

**Spacing:** Single or 1.5 Spacing

**Document Margins:** 2.5cm for top, bottom, left and right.

### **Use of Generative AI:**

If you are using Generative AI (e.g. ChatGPT, etc.) as a tool to assist you in completing the assignment, please do the following:

- Cite the tool for in-text citation and also under reference list. You can refer to the following RMIT Library websites for referencing guidelines:  
[https://rmit.libguides.com/referencing\\_AI\\_tools/home](https://rmit.libguides.com/referencing_AI_tools/home)  
[https://rmit.libguides.com/referencing\\_AI\\_tools/images](https://rmit.libguides.com/referencing_AI_tools/images)

- Compile all evidence of using Generative AI by providing the screenshots to show how you prompted the tool and what responses you received under Appendix.

**Some notes and guidelines (in addition to Lecture 12 slides):**

- Include and discuss all the questions given in *Part 2) Questions from the new client* in your report.
- Do not need to access MySQL server to work on *Question 2 SQL Scripts*. You can make use of tables with sample data (or examples of records) to support your explanations. The tables could be created by typing in MS Word or MS Excel.
- Use a diagramming tool(s) to produce your diagrams for *Question 1: Data Modelling*. Do not submit diagrams that are drawn using pens/ pencils and papers.
- Organise the report with sections and subsections.
- Include justifications and explanations of the SQL queries in your report.
- Consider that the report covers a wider audience, including management and business users as well as developers.
- Label the print screens, figures, diagrams, and tables in the report properly.
- Include References and Appendix sections in your report as you see fit.
- Use Harvard referencing style if you are citing references from web resources to support your discussions.

**Submission Format:**

For each individual, the following documents must be submitted via Canvas. You can merge the following files and submit as a zip file.

- a single copy of the final report (PDF version)
- Assignment Cover Sheet
- a Turnitin generated report (PDF version or receipt/ evidence of submission to Turnitin if it takes too long to generate the report)
- Orange related zip files (include Orange loadable dataset(s) in .xlsx, Orange workflow(s) in .ows)

You can go to *Assignments | Assessment Task 3: Individual Final Assessment | Assessment Task 3 – Individual Assignment* and click on *Submit Assignment*.

**Use of Turnitin:**

Please note that you need to use Turnitin to self-check your report for compliance of academic integrity and plagiarism detection. Report that is not checked by Turnitin will not be marked.

Turnitin is *not* the submission link of your final report (PDF version). After you have self-checked your report, you need to submit the final version of your report together with its Turnitin report (or receipt/ evidence of submission to Turnitin if it takes too long to generate the report) via the designated submission link.

To self-check your report, you can go to *Assignments | Important – Other Assignment Related Info. | Turnitin - Self-Check Your Report \*Not Final Submission\** and click on *Load Turnitin - Self-Check Your Report \*Not Final Submission\** in a new window.

## 4) Marking Rubric:

### Note:

- 1) Marking rubric shows the mark out of 100 points and the total mark will then be converted to 50%.
- 2) Assurance of Learning (AoL) is included for the purpose of AACSB accreditation process.

Assessed Components	Exceptional	Very Good	Acceptable	Needs Improvement	Poor
#AoL5 Expert: Application & Enhancement of Discipline Practices// <b>Data Modelling with ERD and Relational Model (30 points)</b>	Outstanding demonstration of using appropriate data modelling techniques. Diagrams produced are outstanding, elegant and have very few or no mistakes. (24.00 – 30.00 points)// AoL5: Applies advanced disciplinary techniques and practices to push boundaries and contribute actionable methods for real-world scenarios.	Sound demonstration of using appropriate data modelling techniques. Diagrams produced are sound and have few mistakes. (21.00 – 23.99 points)// AoL5: Analyses advanced disciplinary techniques and practices to develop actionable methods for real-world scenarios.	Satisfactory demonstration of using appropriate data modelling techniques. Diagrams are produced at satisfactory level but still have some mistakes. (18.00 – 20.99 points)// AoL5: Combines disciplinary techniques and practices to contribute to actionable methods for real-world scenarios.	Limited demonstration of using appropriate data modelling techniques. Diagrams are somewhat complete but still have a lot of mistakes. (15.00 – 17.99 points)// AoL5: Matches disciplinary techniques and practices to real-world scenarios.	No demonstration or very little demonstration of using appropriate data modelling techniques. Diagrams are incomplete and have a lot of mistakes. (0.00 – 14.99 points)// AoL5: Unsuccessfully matches disciplinary techniques and practices to real-world scenarios.
#AoL4 Digitally Adept: Translates and conveys technical concepts to varied audiences// <b>SQL Scripts (15 points)</b>	SQL scripts are sophisticated and logically well structured, show great complexity, use appropriate syntax, contain no error and are able to	SQL scripts are logically well structured, show good complexity, use correct syntax but still contain very little errors and can address most/ all	SQL scripts use appropriate syntax, show some complexity but contain reasonable number of errors and can address some business purposes.	SQL scripts somehow use appropriate syntax, show minimal complexity but full of errors and can address limited business purposes. Discussions/ justifications	SQL scripts do not use appropriate syntax, overly simplistic and are unable to address any/ very limited business purpose. Discussions/ justifications are missing or very limited.

	address all business purposes. Discussions/ justifications are critical, well-articulated and rigorous. (12.00 – 15.00 points)// AoL4: Explains complex technical concepts by breaking them down to a basic form to facilitate understanding and their application to varied audiences.	business purpose. Discussions / justifications are sound. (10.50 – 11.99 points)// AoL4: Describes technical aspects and their application to varied audiences.	Discussions / justifications are sufficient. (9.00 – 10.49 points)// AoL4: Outlines technical aspects to varied audiences.	are limited or insufficient. (7.50 – 8.99 points)// AoL4: Defines technical concepts to varied audiences.	(0.00 – 7.49 points)// AoL4: Fails to convey technical concepts.
<b>Data Analytics with Orange (15 points)</b>	Excellent model chosen to produce accurate prediction and visualisation. Discussions/ justifications are insightful, critical, well-articulated and rigorous. (12.00 – 15.00 points)	Good model chosen to produce accurate prediction and visualisation . Discussions / justifications are insightful and sound. (10.50 – 11.99 points)	Appropriate model chosen to produce satisfactory prediction and visualisation . Discussions / justifications are somewhat insightful and sufficient. (9.00 – 10.49 points)	Appropriate model chosen to produce sensible prediction and visualisation. Discussions/ justifications are limited or insufficient and barely have insights. (7.50 – 8.99 points)	Inappropriate model chosen and unable to produce sensible prediction and visualisation. Discussions/ justifications are missing or very limited and have no insights. (0.00 – 7.49 points)
<b>Database Administration Approaches (20 points)</b>	Convincing, critical, well-articulated and rigorous discussions/ justifications for the required recommendations.	Sound discussions/ justifications for the required recommendations. (14.00 – 15.99 points)	Satisfactory discussions/ justifications for the required recommendations. (12.00 – 13.99 points)	Limited or insufficient discussions/ justifications for the required recommendations. (10.00 – 11.99 points)	No or very limited discussions/ justifications for the required recommendations. (0.00 – 9.99 points)

	(16.00 – 20.00 points)				
<b>Big Data and Analytics (20 points)</b>	Convincing, critical, well-articulated and rigorous discussions/justifications for the required suggestions. (16.00 – 20.00 points)	Sound discussions/justifications for the required suggestions . (14.00 – 15.99 points)	Satisfactory discussions/justifications for the required suggestions . (12.00 – 13.99 points)	Limited or insufficient discussions/justifications for the required suggestions. (10.00 – 11.99 points)	No or very limited discussions/justifications for the required suggestions. (0.00 – 9.99 points)
<b>Overall</b>	<b>High Distinction (HD)</b>	<b>Distinction (D)</b>	<b>Credit (C)</b>	<b>Pass (P)</b>	<b>Fail (N)</b>
	<b>80 or more</b>	<b>70 or more</b>	<b>60 or more</b>	<b>50 or more</b>	<b>less than 50</b>