School of Computing and information Systems

# PROGRAMME: BSC. BUSINESS INTELLIGENCE AND DATA ANALYTICS

## BI205 DATA WAREHOUSING Year 2 Semester 2

## ASSIGNMENT + RUBRIC STRUCTURE + SCRIPT

**Hand out Date:** 25 Mar 2022 **Hand in Date:** 25 May 2022

**Total Marks: 100**

### Instructions to candidates

1. Candidates must attempt ALL questions.
2. You are to make your submission on turn-it-in. You may consult with your tutor/lecturer on how this will be done.
3. Your Assignment submission must have a cover page with full student details. The cover page will be provided to students when the Assignment is issued out. On the cover page, you will find an acknowledgement statement which must be signed by the student as proof or admission or affirmation of one own's work being submitted.
4. Ensure that you have an account on turn-it-in by going to www.turnitin.com. Use the credentials provided for your account, for accessing this system. If you do not have your turn-it-in account credentials get hold of the module tutor/lecturer as soon as possible.
5. If there is program code to be submitted, ensure that your folder has been created by your lecturer on the submission Server i.e., Sechaba, and you are able to submit inside the folder.
6. Any work with a plagiarism level above **30 % will not be marked.** It is your responsibility to make sure that the plagiarism level detected in your work is within this level. Monitor the plagiarism rating of your work on regular bases. If you share your solution with others, chances of the plagiarism rising above this level are high.
7. **It is your responsibility to ensure that the Introduction to Data Analysis module is in turn-it-in and that you able to see it before the submission date, so you can submit your report on the module link/bin. Consult with your tutor/lecturer if this is not the case.**
8. **Save the file name using the following convention or format surname_firstname_cohort_assignment code.docx or .pdf e.g., smith_david_april2019_A03.docx**
9. **Note that this assignment may be subject to change or amendment and that care shall be taken to ensure that any such amendment or change shall not prejudice or disadvantage you/the candidate/the student.**

**This Assignment consists of NINE (9) printed pages excluding the cover page**

Appendix 1: Assignment submission cover sheet

**ASSIGNMENT SUBMISSION COVER SHEET**

| | |
|---|---|
| **Student Id:** | |
| **Student names:** | |
| **Student email**: | |
| **Cohort**: | |
| **Assignment title**: | |
| **Date of submission**: | |
| **Programme of Study**: | |
| **Year of Study**: | |

**Intellectual property statement**

By checking the box below, I certify that this assignment is my own work and is free from plagiarism. I understand that the assignment may be checked for plagiarism by electronic or other means and may be transferred and stored in a database for the purposes of data-matching to help detect plagiarism. The assignment has not previously been submitted for assessment in any other unit or to any other institution.

**I have read and understood the Botswana Accountancy College plagiarism guidelines policy.**

☐ Agree

**Signature :** ……

**Date :** ……..

# Assignment Learning Outcomes :

**Students should be able to:**

*BIS205-LO1 – Demonstrate understanding of concepts and architectures of data warehousing*

*BIS205-LO3 - Gather requirements for data warehousing*

*BIS205-LO4 - Explain data warehouse operations*

*BIS205-LO5 - Design a dimensional model for data warehousing*

*BIS205-LO6 - Design a physical model for data warehousing*

*BIS205-LO7 - Comprehend extract, transform and load (ETL) strategies*

*BIS205-LO11 - create a data warehouse with Oracle or SQL Server DBMS,*

# Assignment Title : Vehicle Sales Data Warehousing

## The assignment is made up of two parts as follows :

**Part A: Data Warehouse Solution Documentation Report (TurnItIn Submission) : 50%**

**Part B: Data Warehouse Solution, Design and Implementation using Oracle or MS SQL Server DBMSs (Practical To Present Solution) : 50%**

This assignment will attract a zero mark if plagiarism percentage score is above 30 after being submitted on Turnitin.

Please consult with your client who is your tutor / lecturer throughout the assignment where you need assistance and / or clarification of the requirements or tasks.

# Scenario- AZ-Motors Data Warehousing

AZ-Motors is an international organization specializing in car sales. The organization has been established in 1982 and has been in this business for the past 40 years and are now masters of their trade. They use a Relational Database Management System (RDBMS) to manage their customers, products, staff, and sales details. This Online Analytical Processing (OLTP) system is serving them effectively well for their daily transactions.

The organization has amassed large volumes of data over their operational years. The management has learned about Business Intelligence and got interested since they have large volumes of transactional data. They have approached you as a specialist in Business Intelligence (BI) to assist with the implementation of a Data Warehouse (DW) which will help them to effectively use their historical data to produce trending, patterns, and analytics, that may be used for forecasting future sales based on past and current business activity. This is envisaged to improve the company's strategic intent, focus and market share as the intelligence analytics will provide red flags to alert management of areas of improvement to avert potential losses, and green flags to alert managers of where they are doing right to further strengthen their market share.

The following Entity Relationship Diagram (ERD) represents the OLTP database (source system) for AZ-Motors, and the associated database script is provided. Analyse the ERD and assist them in their quest for excellency through BI by completing the following tasks;

# AZ-Motors OLTP Database : CarSales_DB

| VEHICLE | |
|---|---|
| VEHICLEID | Int (PK) |
| MAKEID | Int |
| MODELID | Int |
| COSTPRICE | Decimal(18,2) |
| PRICE | Decimal(18,2) |
| QUANTITY | Int |
| DATEMANUFACTURED | Date |

| MAKE | |
|---|---|
| MAKEID | int |
| MAKE | NVarchar(50) |

| MODEL | |
|---|---|
| MODELID | int |
| MODEL | NVarchar(50) |
| BODYSTYLE | NVarchar(50) |
| MAKEID | int |

| CUSTOMER | |
|---|---|
| CUSTOMERID | Numeric (PK) |
| FIRSTNAME | NVarchar(30) |
| SURNAME | NVarchar(30) |
| GENDER | NVarchar(1) |
| DOB | Date |
| CELLNO | NVarchar(20) |
| ADDRESS | NVarchar(250) |
| EMAIL | NVarchar(100) |
| COUNTRYID | Int |
| DATEREGISTERED | Date |
| STATUSID | Int |

| SALESTRANSACTION | |
|---|---|
| TRANID | Numeric |
| TRANDATE | Date |
| TRANTIME | Time |
| CUSTOMERID | Numeric |
| VEHICLEID | Int |
| PRICE | Decimal(18,2) |
| QUANTITY | Int |
| AMOUNT | Decimal(18,2) |
| BRANCHID | Int |
| STAFFID | Numeric |
| STATUSID | Int |
| STATUSREASON | NVarchar(250) |

| CUTOMERSTATUS | |
|---|---|
| STATUSID | int |
| CUSTOMERSTATUS | NVarchar(50) |
| DESCRIPTION | NVarchar(250) |

| STAFF | |
|---|---|
| STAFFID | Numeric |
| FIRSTNAME | NVarchar(30) |
| SURNAME | NVarchar(30) |
| GENDER | NVarchar(1) |
| DOB | Date |
| CELLNO | NVarchar(20) |
| ADDRESS | NVarchar(250) |
| EMAIL | NVarchar(100) |
| BRANCHID | Int |
| POSITIONID | int |
| DATEJOINED | Date |
| STATUSID | Int |

| TRANSACTIONSTATUS | |
|---|---|
| STATUSID | int |
| TRANSACTIONSTATUS | NVarchar(50) |
| DESCRIPTION | NVarchar(250) |

| BRANCH | |
|---|---|
| BRANCHID | int |
| BRANCHNAME | NVarchar(50) |
| TELNO | NVarchar(50) |
| ADDRESS | NVarchar(250) |
| EMAIL | NVarchar(100) |
| COUNTRYID | Int |

| STAFFSTATUS | |
|---|---|
| STATUSID | int |
| STAFFSTATUS | NVarchar(50) |
| DESCRIPTION | NVarchar(250) |

| POSITION | |
|---|---|
| POSITIONID | int |
| POSITIONNAME | NVarchar(50) |
| DESCRIPTION | NVarchar(250) |

| COUNTRY | |
|---|---|
| COUNTRYID | Int |
| COUNTRY | NVarchar(100) |
| CAPITALCITY | NVarchar(100) |
| CONTINENT | NVarchar(50) |
| CONTINENTID | Int |
| ISO2_CODE | NVarchar(10) |
| ISO3_CODE | NVarchar(10) |
| IOC_CODE | NVarchar(10) |
| TLD | NVarchar(10) |
| CURRENCY | NVarchar(50) |
| TEL_CODE | NVarchar(10) |
| ICO | IMAGE |
| PREFERENCE_ORDER_ID | Int |

| CONTINENT | |
|---|---|
| CONTINENTID | Int |
| CONTINENT | NVarchar(50) |
| AREA_KM2 | Numeric(18,0) |
| POPULATION | Numeric(18,0) |
| MOST_POPULOUS_CITY | NVarchar(100) |
| CODES | Int |

NB: The black broken relationship arrows from CUSTOMER and BRANCH tables connect to the arrow in the COUNTRY representing a Foreign Key relationship.

## You are provided with the SCRIPT to generate this database

## Part A: Assignment Total Marks [50 marks]

**Write a short report explaining the following tasks in your DW solution.**

1. **Formulate the Strategic Questions for the Data Warehouse [5 Marks]**
   Formulate 5 (five) strategic questions that will assist in coming up with relevant dimensions as the first step in the process of creating a Data Warehouse.

2. **Create a Data Dictionary [20 Marks]**
   Come up with the dimensions for each strategic question in 2 above, specifying the relevant dimension names, column names, datatypes and field widths in the following format;

| **DIM_NAME** | | |
|---|---|---|
| COLUMNNAME1 | DATATYPE | PK |
| COLUMNNAME2 | DATATYPE (N) | |
| COLUMNNAME3 | DATATYPE (N) | FK: References DIM_NAME(COLUMNNAME) |

   Create a FACT Table for storing the analytics for your DW Solution using the same format above

3. **Create a Star Schema [5 Marks]**
   Produce a star schema equivalent of the E/R model for the AZ-Motors scenario, using the information in the Data Dictionary in 3 above.

4. **Extraction, Transformation, and Loading (ETL) Process**
   Explain the Extraction Transformation and Loading (ETL) Processes you are going to implement in your DW solution. **[15 Marks]**

5. **Analytic Queries and Analytics for Decision Making**
   Explain your analytic queries, how they work and what form of analytics will be generated and how they will assist the organization in turning around its fortunes. **[5 Marks]**

## Part B: Assignment Total Marks [50 marks]

1. **Run provided Script to create the OLTP Database with ERD above [5 Marks]**
   Execute the script to generate the OLTP database for this scenario.

2. **Create your DW Solution from your Star Schema [5 Marks]**
   Create the Data Warehouse for the AZ-Motors based on your Star Schema in Part A above.

3. **Extract, Transform, and Load Data into the Data Warehouse [15 Marks]**
   Write SQL Scripts, or procedures, or packages to assist in extracting, transforming and loading data into the DW.

4. **Create Analytic Queries [15 Marks]**
   Write 3 (three) analytic queries to mine and generate intelligence from the DW which can be used for decision making to turn around the fortunes of the company.

5. **Create Reports [5 Marks]**
   Create 1 (one) BI report using one of your analytic queries in 6 above. Reports may be created using SQL Server Reporting Services (SSRS), Oracle Reporting, HTML, Power BI, Python Libraries, Jasper, C# or any technology you are familiar with.

6. **Solution Demonstration / Presentation to Client [5 Marks]**
   Conduct a demonstration / presentation of the solution to your client to showcase the capabilities of the solution and convince the client that BI is the way to go.

# MARKING RUBRIC

## Part A: Rubric Total Marks [50 marks]

1. **Formulate the Strategic Questions for the Data Warehouse [5 Marks]**
   - 1 mark for each meaningful strategic question formulated, max 5 , total 5 marks

2. **Create a Data Dictionary [20 Marks]**
   For each dimension and Fact Table created
   - 1 mark for each meaningful dimension name
   - 1 mark for correct attributes
   - 1 mark for correct datatype and field size
   - 1 mark for any correct Primary and Foreign key identified
   - Total marks 4 per dimension/fact table x 5 = 20 Marks

3. **Create a Star Schema [5 Marks]**
   - 1 mark for each correct relationship represented in the star schema, max 5 relationships considered (5 marks)

4. **The Extraction Transformation and Loading (ETL) Processes [15 Marks]**
   - 5 marks for correct performed Extraction process explanation
   - 5 marks for correct performed Transformation process explanation
   - 5 marks for correct performed Loading process explanation

5. **Analytic Queries [5 Marks]**
   - 2 marks for the explanation of the analytic queries created
   - 3 marks for the analytics generated and explanation of the relevance and usage of the analytics in assisting the organization to improve their business.

## Part B: Rubric Total Marks [50 marks]

1. **Run provided Script to create the OLTP Database with ERD above [5 Marks]**
   - 5 marks for successfully running the script to create the OLTP database
2. **Create your DW Solution from your Star Schema [5 Marks]**
   - 1 mark for each correctly created dimension / fact table, max 5

3. **Extract, Transform, and Load Data into the Data Warehouse [15 Marks]**
   - 1 mark for each correct data extraction SQL statements, max 5, (5 marks)
   - 1 mark for each data transformation operation performed, max 5, (5 marks)
   - 1 mark for each correct data loading SQL statement/operation, max 5, (5 marks)

4. **Create Analytic Queries [15 Marks]**
   - For each of the 3 (three) required analytic queries
   - 1 mark for each aggregate function correctly used max 2, (2 marks)
   - 1 mark for showing grand totals
   - 2 marks for the correct and meaningful results displayed

5. **Create Reports [5 Marks]**
   - 2 marks for proper report design using an appropriate tool
   - 2 marks for correctly displaying data
   - 1 mark for user friendliness

6. **Solution Demonstration / Presentation to Client [5 Marks]**
   - 3 marks for demonstration of solution functionality
   - 2 marks for confidence of presentation and showing ownership of work

-------- **END OF ASSIGNMENT** --------