# COMP3350

# Assignment 2 – Business Intelligence

### Semester 1, 2022

**Due Date** Apr 26th 10am

Assignment 2 is due on Apr 26th 10am. Each group will

- upload the assignment files to Canvas and
- present your BI report and demonstrate your assignment on tutorial session on Apr 26th

All members must be present for the demonstration. Groups without demonstrations will be penalised for sections which are not demonstrated.

## Weighting

20% of course mark

## Assignment Information

This is a group assignment. This assignment has 3 sections.

## Group Formation

You need to work in groups of two for this assignment. You may continue with your group members in Assignment 1 or form different groups for Assignment 2**.** If you are forming a different group from Assignment 2, you must email the lecturer the group formation by the tutorial session on 5th April. Otherwise, your group is assumed to be the same as in Assignment 1.

## Assignment Specifications

This assignment has 3 sections.

### Section 1: Datawarehouse Design (5 marks)

In this section you will design a data warehouse schema for the University (the scenario discussed in Assignment 1).

Design a data warehouse schema to satisfy University's decision makers' information needs. You need to only design the data warehouse **schema only**. You do not need to implement it.

You need to write a short report explaining subject-area/s covered by your data warehouse, illustrate the documented schema and discuss how the data warehouse

satisfies the information analysis needs of the University. Give examples of analysis queries that your design supports.

Save your document as *DatawarehouseDesign_University _<team number>.docx*.

**Section 2: ETL Exercise (5 marks)**

In this section, you will create an ETL task to load Customer Data to a table. Your group is already provided with sample text data of Customer data.

- Customer data: Data on customers (CustomerData.txt)

Your group is asked to load this data into SQL Server database called *Staging_Area* by creating a Server Integration Services (SSIS) project, called *Assignment3_<team number>_ETLExercise*

Next, create an SSIS Package called *LoadingCustomerData* to load Customer data.

| Input File | Package Name | Destination Table Name |
|---|---|---|
| CustomerData.txt | LoadingCustomerData.dtsx | CustomerDimension |

Ensure that the following data type conversions are included in the loading package for data:

| Destination Table | Column Name | Data Type |
|---|---|---|
| CustomerDimension | CustomerKey | Integer |
| | DateOfBirth | Date |

You have been informed that there have been errors in your SSIS package when loading Customer data. That is, in some records, the following fields have incorrect data:

- City
- StateProvinceCode
- StateProvinceName
- CountryCode
- CountryName
- PostCode

You need to create a SSIS package called *LoadingCustomersCorrected* which loads the data correctly for all records in the *CustomerData.txt* file

Hints:

- In *CustomerData.txt* file, the *StreetAddress* field in some instances contains a "," (comma) which is also the value used to denote the end of a field value.
- You may need to explore constructs for data cleaning and transformation in SSIS such as Conditional Splits and Derived Columns*

## Section 3: Business Intelligence Report (10 marks)

Download and restore the WorldWideImporters (WWI) Data Warehouse sample database. Download *WideWorldImportersDW-Full.bak* file from https://github.com/Microsoft/sql-server-samples/releases/tag/wide-world-importers-v1.0.

The following information about WWI are extracted from https://docs.microsoft.com/en-us/sql/samples/wide-world-importers-what-is?view=sql-server-2017

*Wide World Importers (WWI) is a wholesale novelty goods importer and distributor operating from the San Francisco bay area.*

*As a wholesaler, WWI's customers are mostly companies who resell to individuals. WWI sells to retail customers across the United States including specialty stores, supermarkets, computing stores, tourist attraction shops, and some individuals. WWI also sells to other wholesalers via a network of agents who promote the products on WWI's behalf. While all of WWI's customers are currently based in the United States, the company is intending to push for expansion into other countries.*

*WWI buys goods from suppliers including novelty and toy manufacturers, and other novelty wholesalers. They stock the goods in their WWI warehouse and reorder from suppliers as needed to fulfil customer orders. They also purchase large volumes of packaging materials, and sell these in smaller quantities as a convenience for the customers.*

*Recently WWI started to sell a variety of edible novelties such as chilli chocolates. The company previously did not have to handle chilled items. Now, to meet food handling requirements, they must monitor the temperature in their chiller room and any of their trucks that have chiller sections.*

### *Workflow for warehouse stock items*

*The typical flow for how items are stocked and distributed is as follows:*

- *WWI creates purchase orders and submits the orders to the suppliers.*
- *Suppliers send the items, WWI receives them and stocks them in their warehouse.*
- *Customers order items from WWI*
- *WWI fills the customer order with stock items in the warehouse, and when they do not have sufficient stock, they order the additional stock from the suppliers.*
- *Some customers do not want to wait for items that are not in stock. If they order say five different stock items, and four are available, they want to receive the four items and backorder the remaining item. The item would them be sent later in a separate shipment.*
- *WWI invoices customers for the stock items, typically by converting the order to an invoice.*

- *Customers might order items that are not in stock. These items are backordered.*
- *WWI delivers stock items to customers either via their own delivery vans, or via other couriers or freight methods.*
- *Customers pay invoices to WWI.*
- *Periodically, WWI pays suppliers for items that were on purchase orders. This is often sometime after they have received the goods.*

### *Additional workflows*

*These are additional workflows.*

- *WWI issues credit notes when a customer does not receive the good for some reason, or when the goods are faulty. These are treated as negative invoices.*
- *WWI periodically counts the on-hand quantities of stock items to ensure that the stock quantities shown as available on their system are accurate. (The process of doing this is called a stocktake).*
- *Cold room temperatures. Perishable goods are stored in refrigerated rooms. Sensor data from these rooms is ingested into the database for monitoring and analytics purposes.*
- *Vehicle location tracking. Vehicles that transport goods for WWI include sensors that track the location. This location is again ingested into the database for monitoring and further analytics.*

Part A: Reporting (2 marks)

You are asked to create a report using SQL Server's Reporting Services. Save the project as *Assignment2_<team number>_SQLReports*. The report provides the monthly and yearly sales for **California** for all years.

The format of the report is given below:

**Monthly-Yearly Sales Report**
**(California)**

| Year | Month | Monthly Sales Amount |
|---|---|---|
| 2013 | January | $2,345.98 |
| | February | $2,532.99 |
| … | | |
| | **Yearly Sales** | **$1,232,322.99** |
| 2014 | **…** | |

Note that the data is sample data and does not pertain to correct values in the actual database.

Part B: Data Analytics (8 marks)

Explore the data warehouse schema and data. Select subject area(s) that your group would like to analyse in WWI. Create data mart(s) either using SQL Server Analysis Server. Save the project/file as *Assignment2_<team number>_DataMarts*.

Analyse the data and write a Business Intelligence report based on World Wide Importers data. Use Excel's BI features such as pivot tables, charts and graphs; and/or PowerBI's visuals, dashboards etc. in your analysis, data visualisation and presentation.

Note that your BI report is presented to the business management team of World Wide Importers such as CEO and senior management, so your BI report should be understood by business decision makers of WWI.

Write a report detailing data analysis, information discovered and present helpful insights and actions items from your data analysis. Use appropriate tables, charts, graphs etc. to present your findings. In addition to the written report, you need to present your BI report to class on Apr 26th tutorial session. Your group's presentation should not exceed 10 minutes.

Save your report as *BusinessIntelligenceReport_<team number>.docx*.

**Submission**

Your submission to this assignment contains 3 parts:

**Section 1:** A softcopy of *DatawarehouseDesign_University_<team number>.pdf* document with a signed group assessment cover sheet submitted via *Assignment2_Section1* link.

**Section 2:** Your root assignment folder zipped named as *Assignment2_<team number>.zip* and submitted to Canvas link.

**Section 3:** A softcopy *BusinessIntelligenceReport_<team number>.pdf* should be submitted to *Assignment2_Section3* link.

The root folder *Assignment2_<team number>* should contain the *Setup.docx* file which outlines any specification for installation and configuration for the submitted project and the following *subfolders*:

| Subfolder | Contents | Description |
|---|---|---|
| *Assignment2_<team number>_DatawarehouseDesign* | *DatawarehouseDesign_ University_<team number>.pdf* | This subfolder contains Data Mart Design document for in section 1. |
| Data | CustomerData.txt | The data files |
| *Assignment2_<team number>_ETLExercise* | SSIS project files | This subfolder contains all SSIS project files for section 2 |
| *Assignment2_<team number>_SQLReports* | SQL Server Reporting Services files | This subfolder contains SSRS project files used in section 3 |
| *Assignment2_<team number>_DataMarts* | SQL Server Analysis project files | This subfolder contains SQL Server Analysis Services project files or Excel file used in section 3 |
| *Assignment2_<team number>_BIReport* | PDF document of BI report, Excel files used to generated reports and presentation slides of the presentation. | This subfolder contains BI report and any Excel/PowerBI files used for data analysis in section 3. |

A **group demonstration and presentation** will be held by the tutor on Apr 26th tutorial session. Each group member **must** be present to explain the implementation of the Assignment. Projects that are not demonstrated and presented will not be marked and may result in a zero mark for the Assignment.

The assessment RUBRIC is given below:

| | Excellent | Good | Satisfactory | Poor | Fail |
|---|---|---|---|---|---|
| Datawarehouse Design (5) | (5) | (4) | (3) | (1-2) | (0) |
| | The data warehouse design is illustrated, documented and clearly justified. The team demonstrates fluency in design and demonstrates fluency at the demo. | The design meets most of the requirements specified. The schema is well documented and justified. The team demonstrates fluency in the provided design. | A schema for data mart is given. The schema is partially documented. | An undocumented and poorly justified data mart schema is provided. | No submission |
| Loading Data (5) | (5) | (4) | (3) | (1-2) | (0) |
| | SSIS packages created and executed without any errors. All members of the group demonstrate fluency. | SSIS packages are created. Demonstrated fluency and execution. Minor errors. | 1 SSIS package created and loading data. 2nd SSIS package has major errors. | At least 1 SSIS package is created and running to load data. | No SSIS packages created |
| SQL Report (2) | (2) | | (1) | | (0) |
| | Reports generated without any errors. The formatting is as specified in the specification. Fluency demonstrated by each member | | Report contains most data correctly. Minor errors in the formatting. Partial fluency in report design and data generation for reports. | | No Reports submitted |
| Data Analytics & BI Report (8) | (7-8) | (5-6) | (3-4) | (1-2) | (0) |
| | Data Mart(s) design and implementation without flaws. A well written report with detailed data analysis supporting conclusions/claims and is fluently presented. Demonstrates fluency in design, implementation, data analysis and presentation | Data Mart(s) design and implementation is presented. A report with data analysis supporting conclusions/claims is presented. Demonstrates fluency in design, implementation, data analysis and presentation | Data Mart is correctly implemented. May contain minor design flaws (e.g. hierarchies). Basic OLAP reports and simplistic data analysis and evaluation | Data Cube design is correct. Errors in implementation and cannot populate cube. No report submitted. | No submission |