INFT3007 - The Information Resource

Assignment 3 – Data Warehouses and Business Intelligence

Semester 1, 2017

Due Date May 24th 9:00am

Assignment 3 is due on May 24th 9:00am. Each group will

- upload the assignment files to the Blackboard and
- submit 1 hard copy of the assignment with a signed coversheet during tutorial on May 24th
- present your BI report and demonstrate your assignment on Friday (May 26th) tutorial session

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 for this assignment. In exceptional cases, a group of one is allowed. You may continue with your group members in Assignment 2 or form different groups for Assignment 3**.** If you are forming a different group from Assignment 2, you must hand in the group formation by the tutorial session on 5th May. Otherwise, your group is assumed to be the same as in Assignment 2.

Assignment Specifications

This assignment has 3 sections.

Section 1: Datawarehouse Design (5 marks)
In this section you will design a data warehouse schema for *Office Wizard* (the scenario discussed in Assignment 1).

Design a data warehouse schema to satisfy *Office Wizard*'s decision makers' information needs and providing a consolidated enterprise view of the data. 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-areas covered by your data warehouse, illustrate the documented schema and discuss how the data warehouse satisfies the information analysis needs of *Office Wizard*. Give examples of analysis queries that your design supports.

Save your document as *DatawarehouseDesign_OfficeWizard_<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

Find the error and create a new 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 AdventureWorks Data Warehouse database. Adventure Works is a fictitious company that manufactures and sells metal and composite bicycles to North American, European and Asian commercial markets. Adventure Works Data Warehouse contains a number of data marts for the fictitious Adventure Works company.

Part A: Reporting (2 marks)

Create the following report using SQL Server's Reporting Services. Save the project as *Assignment3_<team number>_SQLReports*. The report provides the monthly and yearly reseller sales for **European territory** for all years.

The format of the report is given below:

Monthly-Yearly Sales Report (Australia)

Year	Month	Monthly Sales Amount
2001	January	$2,345.98
	February	$2,532.99
…		
	Yearly Sales	**$1,232,322.99**
2002	**…**	

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 would like to analyse in Adventure Works. You can either create

(i.) a multidimensional in SQL Server Analysis Services (SSAS) or

(ii.) a PowerBI data model

for data analysis. Save the project/file as *Assignment3_<team number>_DataMart*.

Analyse the data and create a Business Intelligence report based on Adventure Works data. Use data visualisations such as charts, graphs, dashboards etc. in your analysis, data visualisation and presentation.

Note that your BI report is presented to the business management team of Adventure Works such as CEO and senior management, so your BI report should be understood by business managers.

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

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

Progress Monitoring

In this assignment, progress will be monitored for each group at tutorials. Please ensure that all group members attend the tutorial sessions to show their group's progress.

Sections	**Tutorial Session**
Sections 1 & 2	May 12th Tutorial Session
Section 3	May 19th Tutorial Session

Submission

You submission to this assignment contains 3 parts:

Section 1: A hardcopy of *DatawarehouseDesign_OfficeWizard_<team number>.docx* document with a signed group assessment cover sheet and softcopy submitted to *Assignment3_Section1* TurnItIn link.

Section 2: Your root assignment folder zipped named as *Assignment3_<team number>.zip* and submitted to Blackboard.

Section 3: A hardcopy of *BusinessIntelligenceReport_<team number>.docx* with a group assessment cover sheet and a group presentation on May 26th tutorial session. A softcopy *BusinessIntelligenceReport_<team number>.docx* should be submitted to *Assignment3_Section3* TurnItIn link.

The root folder *Assignment3_<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**
Assignment3_<team number>_DatawarehouseDesign	Word document *DatawarehouseDesign_OfficeWizard_<team number>.docx*	This subfolder contains Data Warehouse Design document for in section 1.
Data	CustomerData.txt	The data files
Assignment3_<team number>_ETLExercise	SSIS project files	This subfolder contains all SSIS project files for section 2
Assignment3_<team number>_SQLReports	SQL Server Reporting Services files	This subfolder contains SSRS project files used in section 3 PART 1
Assignment3_<team number>_DataMarts	SQL Server Analysis project files or PowerPivot Excel file	This subfolder contains SQL Server Analysis Services project files or Excel file used in section 3

Assignment3_<team number>_BIReport	Word document of BI report, Excel files used to generated reports and presentation slides of the presentation.	This subfolder contains BI report and any Excel files used for data analysis in section 3.

A **group demonstration and presentation** will be held by the tutor on **May 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.

References

* SSIS help topics and videos can be found at https://code.msdn.microsoft.com/site/search?f%5B0%5D.Type=User&f%5B0%5D.Value=SQL%20Server%20Integration%20Services&pageIndex=1

https://code.msdn.microsoft.com/site/search?f%5B0%5D.Type=User&f%5B0%5D.Value=SQL%20Server%20Integration%20Services&pageIndex=2

***** PowerBI Guided Learning - https://powerbi.microsoft.com/en-us/guided-learning/

The assessment RUBRIC is given below:

	Excellent	Good	Satisfactory	Poor	Fail
Data Warehouse 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 warehouse is given. The schema is partially documented.	An undocumented and poorly justified data warehouse 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)	(5)	(4)	(3)	(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

Group Formation Form

GROUP NO: _____ (TO BE FILLED BY LECTURER)

GROUP MEMBERS:

I agree to participate in the mentioned group for Assignment 3 of INFT3007

Student ID	**Name**	**Signature**