Assignment 4 (10%)
Date Given: Mar 27, 2019
Submission Due: Apr 10, 2019 at 11:59 pm (midnight)
**** Late submissions are not accepted and will result in a 0 on the assignment**

Objective:

This assignment covers concepts related to Business Intelligence. Consider this assignment as the Analysis phase of an industry project. The designed BI framework in this assignment will answer all business questions for a specific client.

Grading Scheme:

- Data Gather/Upload: 10%
- R1: Background research on Fact & Dimension tables: 10%
- R2: Report containing Fact & Dimension tables for Parks, Crime, Business, and the answers with screenshots: 20%
- Parks BI implementation: 20%
- Crime BI implementation: 20%
- Business BI implementation: 20%

Academic Integrity:

- This assignment does not require group work. Therefore, each student is expected to complete their work by themselves. Collaboration of any type amounts to a violation of the academic integrity policy and will be reported to the AIO.
- Do not copy texts verbatim from online or printed materials
- Do not copy texts from other's work
- Do not submit other's work
- If you obtain help from Tutor(s), please acknowledge
- Provide citation for texts, images, tables, data etc. **Follow IEEE/APA/MLA citation styles**
- The Dalhousie Academic Integrity policy applies to all material submitted as part of this course. Please understand the policy, which is available at: https://www.dal.ca/dept/university_secretariat/academic-integrity.html

Hypothetical Scenario:

HalifaxInfo is trying to identify key performance indicators (KPIs) in the Halifax region to improve the *business*, *education*, *lifestyle*, and *safety*. In the first phase of the project, the company has gathered relevant structured data and information that are collected by various sources. The company found key entity sets, their attributes, and relationships that are critical to the future system. In the second phase, the company has setup a Big data infrastructure and implemented a parallel computing framework using MapReduce. In the first two phases, the company gathered relational data and unstructured data (tweets) on Halifax and performed some basic study. In addition to sentiment analysis of tweets, *HalifaxInfo* has successfully completed a pilot project (3rd phase) on analysis of some historical news data.

The fourth phase of the project focuses on implementing BI frameworks for parks, crime, and business. The goal of this phase is to extract useful information for decision making. A client of *HalifaxInfo* is going to take major decision about establishing schools, real-estate, shopping malls in the Nova Scotia province. As the information specialist, your job is to build separate BI frameworks for each domain, and then find the correlation between the domains and various factors.

***** Your Tasks for this Assignment *****

As an *information specialist,* you are expected to perform a series of tasks that are mentioned in section A to section D.

Specific Tasks (section A to D) – complete them as specified and submit on Brightspace based on the submission instructions.

A. Data Gathering and Upload:

1. Use your cloud instance(s) or local server to perform this task.
2. Use your assignment 1 data on Parks, Crime, and business.
3. If your park data is not sufficient gather more data from https://data.novascotia.ca/Lands-Forests-and-Wildlife/DNR-Camping-Parks-Reservation-Data-2016/4zt7-x443
4. If your data on Business is insufficient, you might want to visit these links: [a]https://data.novascotia.ca/Business-and-Industry/Nova-Scotia-Co-operatives/k29k-n2db
   [b]https://data.novascotia.ca/Business-and-Industry/Co-operatives-Financial-and-Operating-Summary/ff6i-nhbm
5. Upload the data on a suitable database server. If you have the data on a RDBMS, you can avoid this step.

B. Exploratory Study (R):

6. Study Fact tables and Dimension tables from authentic sources, and write one paragraph (1/2 page) report. The report must reflect your understanding about the topic. (Do not copy verbatim from any online or offline sources). Provide references in MLA/APA/IEEE format.

C. BI Framework (Parks, Business, Crime):

In this section, you need to design 3 separate BI frameworks. If possible, you can join the BI frameworks.

Now, to build the BI framework, you need to perform the following tasks.

7. Define the main facts to be analyzed (Hint: These facts become the source for the design of the fact table)
8. Define and describe appropriate dimensions. (Hint: These dimensions become the source for the design of the dimension tables.)
9. Define the attributes for each of the dimensions
10. Recommend the appropriate attribute hierarchies
11. Implement your data warehouse design, using the star schema or snowflake schema. Include your star or snowflake schema for each of the domains in the report **R2**.
12. Connect your database to Cognos BI and create the report
13. Perform multidimensional data analysis using the concept of OLAP that you have learned in the BI lesson. Use Cognos BI, Powerplay and necessary tools.

D. Answer the following questions using your multidimensional model:

****Questions from the Client (These are sample questions. You can add more based on your Model):**

- How safe is Halifax or NovaScotia? Is there any increase in crime rate in past 5 years (depending on data availability, you can change this more or lesser years/ months)?
- Which region/area/street has less crimes?
- Is time related to crime?

- Which parks in NovaScotia allow Senior rate?
- Is there a growth in park visitors? Past 5 years' data should be analyzed.
- In past 5 years, how many cooperatives have submitted their report? Is there any specific year, which has more submissions?

  Your answers for the above questions must use Drill-down, roll-up approach, slice-dice.

Submission Instruction:

- Create a Folder with your name and B00 number, and store all your files –
  - PDF files of R1(Study), and R2(Model, Answers, Screenshots).
- Compress the folder and create a .ZIP file (do not use other compression formats)
- Upload the .ZIP file on Brightspace.
- Submission Due: **Apr 10, 2019 at 11:59 pm (midnight)**