5408 Assignment 3

Ruize Nie

B00739439

A part Semantic Analysis of Twitter

1. Tweets capture using Tweepy. (show only top 20, rest attach on folder)

A screenshot of a cell phone

Description automatically generated

1. Tweets semantic analysis [1] (show only top 20, rest attach on folder)

A screenshot of a cell phone

Description automatically generated

1. Visualize the most frequently occurring words in the positive and negative tweets

A screenshot of a social media post

Description automatically generated

B part Semantic Analysis of News Article

1. News capture using News API. (show only top 20, rest attach on folder)

A close up of a newspaper

Description automatically generated

1. Term frequency-inverse document frequency

A screenshot of a cell phone

Description automatically generated

1. Highest occurrence of the word “Canada” (show only top 20, rest attach on folder)

A screenshot of a cell phone

Description automatically generated

1. Program output

A close up of a logo

Description automatically generated

C part Business Intelligence of Dalhousie University Data Model

1. For the Dalhousie University data model, the main facts are the sum of the number of programs within each department by times.
2. In the Dalhousie University data model, the possible dimensions which can be used in the dimension table are time, semester, campus & department. Each of the dimension provide source to the total number of programs in the fact table.
3. The attributes of the different dimensions of Dalhousie University data model are as:
   1. Time dimension: time\_id, time\_description, beginning\_time, ending\_time
   2. Semester dimension: semester\_id, semester\_description, semester\_time, ending\_date, begin\_date.
   3. Campus dimension: campus\_id, campus\_description, campus\_locaton
   4. Department dimension: department\_id, department\_description, department\_location
4. For the Dalhousie University data model, the attribute hierarchy from top to down is as:

A close up of a sign

Description automatically generated

1. Data warehouse design

A picture containing screenshot, text

Description automatically generated

1. Connect database with Cognos BI [2]
   1. Check how many programs offer by each department
   2. Only show those departments which offer more than one program

A screenshot of a cell phone

Description automatically generated

1. Multidimensional data analysis
   1. Analysis Number of Department each Campus [2]

A screenshot of a cell phone

Description automatically generated

* 1. Analysis Number of Employee each Department [2]

A screenshot of a cell phone

Description automatically generated

1. Using my BI Framework, I can only answer the question related to Which Department offer highest number of Program. And from the report it is Agriculture Department, and Computer Science are not offer the highest number of programs. However, Since I don’t have enough data related to courses which department offered. So I can give the report for that.

**Reference:**

[1] A. Navlani, “Latent Semantic Analysis using Python,” 2018, <https://www.datacamp.com/community/tutorials/discovering-hidden-topics-python>.

[2] A. Hussain, “Cognos Analytics 11: Smarter Self-Service BI,” 2018, <https://www.royalcyber.com/blog/business-intelligence/cognos-analytics-11-smarter-self-service-bi/>.