University of Ottawa

School of Electrical Engineering and Computer Science

Introduction to Data Science: CSI4142 2019

Project Phase 1

Total marks: 100

Instructions:

1. Submit your answer, in Adobe PDF format, before the deadline.
2. This is a group project.

The City of Ottawa maintains details of traffic accidents for the period 2014 to 2017, as found at the following four open data links:
http://data.ottawa.ca/dataset/collisiondata2014
http://data.ottawa.ca/dataset/collisiondata2015
http://data.ottawa.ca/dataset/collisiondata2016
http://data.ottawa.ca/dataset/collisiondata2017

In addition, Environment Canada maintains information about the **hourly** weather reports across Canada. This data includes a number of stations in Ottawa, and the hourly readings may be retrieved from http://climate.weather.gc.ca/historical_data/search_historic_data_e.html. (Note that the longitude and latitude of all stations are provided, so it would be appropriate to link an accident to the nearest station.)

The major of Ottawa decides to create an integrated data mart in order to study the interplay between the weather and traffic accidents. He is interested in exploring the impact of severe weather on the frequencies and severities of accidents. In addition, his goal is to identify neighborhoods, streets and intersections where accidents often occur. Further, he aims to assess the trends in types and frequencies of accidents over the four years, notably the trends in numbers of fatal injuries. He is also interested in determining whether there are specific days of the week, or times of the day, when accidents are more frequent. Finally, the perceived interplay between special events (e.g. Winterlude or Fall Rhapsody) and the number of accidents in certain neighborhoods are a concern.

Your Task: Create the dimensional model for this data mart. Be sure to declare your grain and to show all the facts and dimensions.