University of Ottawa

School of Electrical Engineering and Computer Science

CSI4142 Fundamentals of Data Science

Project Phase 1: Conceptual Design

Instructions:

Submit your conceptual design using your group locker in the Virtual Campus. Your document should be in PDF format.

Description:

At a recent meeting of the city council of Ottawa, one on the councillors mentioned that she recently read a very interesting article on the use of data science and artificial intelligence (AI) in crime detection and prevention in cities such as London and Denver.

Suppose that, in subsequent discussions, the city council of Ottawa voted to create a crime data mart to study trends in crime in the Ottawa region. Your team is hired to complete a feasibility study and to produce a proof of concept.

Your team subsequently meets with the city council, and notes the following requirements.

The council is interested in looking into seasonal trends and pinpointing locations that may benefit from additional policing. In addition, the goal is to identify trends in the types and frequencies of violations over time. They are interested in exploring the trends in different types of crimes, such as crimes against a person, crimes against a property, as well as specific crimes such as fraud and theft. The council also wants to explore whether there are specific times of the day, or days of the week, when certain crimes occur more frequently. The council further wishes to determine if there is any interplay between holidays and special events (e.g. New Year, Winterlude or Fall Rhapsody) and the frequency and types of crimes committed. Notably, crimes against properties are perceived to be more frequent during July and August, but this perception is based on hear-say, rather than facts. The council also notes that they would be very interested in contrasting the occurrences in crimes in Ottawa to other major cities, such as Toronto, Vancouver or Edmonton.

When studying the data realities, you realise a major drawback: the city has no detailed data about individual crimes in Ottawa! Rather, they only maintains some yearly statistical data, as located at https://open.ottawa.ca/datasets/crime-statistics-incidents-2014-2018 (last accessed 17/02/2020). It follows that the creation of a detailed data mart would not only be expensive, but also involve substantial data collection. In order to win this contract, your team would thus need to convince the city council to secure substantial financial and human resources.

As a proof of concept, your team decides to use some Open Data in your quest to convince City Hall to invest in this project. You agree to start with the data as located at the following two links, which contain to publicly available data obtained from Vancouver and Denver (accessed on 17/01/2020):
Vancouver: https://www.kaggle.com/wosaku/crime-in-vancouver
Denver: https://www.denvergov.org/opendata/dataset/city-and-county-of-denver-crime

Your Task. Create the dimensional model for the crime data mart. Be sure to declare your grain and to show all the facts and dimensions. Note that your data mart should ambitiously "go beyond" the data realities. Your design could benefit from also investigate other Open Data resources.

Notes:

1. See also https://library.carleton.ca/find/gis/geospatial-data/ottawa-police-service-crime-data) for older legacy data (accessed on 17/01/2020).
2. Another interesting link to other open data crime databases and a discussion on machine learning (accessed on 17/01/2020): https://lionbridge.ai/datasets/16-best-crime-datasets-for-machine-learning/