Knowledge Management Systems

OLAP and Visualisation Assignment

(Maximum marks: 35)

(To be completed in groups of 5 or 6, due in class in Week 13)

Please refer to the Cybersecurity data breaches dataset and the description. The dataset refers to security breaches in the United States from 2005 to 2018. You are required to analyse this dataset using OLAP and other tools and prepare a report based on:

1. Preparing a comparative analysis of the number and intensity of breaches by the type of breach over time. Identify any critical trends over time based on the type of business **(5 marks**)

2. Performing a variety of analyses to test the validity of the claim that "..data breaches are becoming more frequent and more severe over time in the US". Use analytics to arrive at inferences that you can show are valid and provide detailed justification for your conclusion. **(10 marks)**

3. Performing a comparative analysis of the number and seriousness of the breaches for the 5 best performing and 5 worst-performing states with regard to security breaches. Please note that you will have to come up with the best definition of 'best performing' and 'worst-performing' based on other indicators. **(10 marks)**

4. Producing comprehensive and informative visual representations/dashboard of the most appropriate data attributes in the dataset that can visually inform security analysts regarding breach type, geographic, temporal trends in security breaches in the United States based on the given dataset. **(10 marks)**

Notes:

- For questions 1, 2, and 3, your answer should include a discussion of the results and the inferences and conclusions that you were able to make.
- You are free to use other relevant datasets to perform your analyses and to support your claims. Make sure you clearly cite data sources and other references you use in the bibliography section of your report.
- You are allowed to use any tool(s) that you feel comfortable with (e.g. Cognos, Tableau, Pentaho, QlikView, etc.).

Legend

Type of Breach

CARD
Fraud Involving Debit and Credit Cards Not Via Hacking (skimming devices at point-of-service terminals, etc.)

HACK
Hacked by an Outside Party or Infected by Malware

INSD
Insider (employee, contractor or customer)

PHYS
Physical (paper documents that are lost, discarded or stolen)

PORT
Portable Device (lost, discarded or stolen laptop, PDA, smartphone, memory stick, CDs, hard drive, data tape, etc.)

STAT
Stationary Computer Loss (lost, inappropriately accessed, discarded or stolen computer or server not designed for mobility)

DISC
Unintended Disclosure Not Involving Hacking, Intentional Breach or Physical Loss (sensitive information posted publicly, mishandled or sent to the wrong party via publishing online, sending in an email, sending in a mailing or sending via fax)

UNKN
Unknown (not enough information about breach to know how exactly the information was exposed)

Type of Business

BSF
Businesses (Financial and Insurance Services)

BSO
Businesses (Other)

BSR
Businesses (Retail/Merchant including Online Retail)

EDU
Educational Institutions

GOV
Government & Military

MED
Healthcare, Medical Providers and Medical Insurance Services

NGO
Non-profits

UNKN
Unknown

Other Data Attributes

Date Made Public

Company

City

State

Total Records

Description of incident

Information Source

Source URL

Year of Breach

Latitude

Longitude